



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY

A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

DETEKCE FIBRILACE SÍNÍ V EKG

EKG BASED ATRIAL FIBRILLATION DETECTION

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Vít Plch

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Marina Ronzhina, Ph.D.

BRNO 2019

Diplomová práce

magisterský navazující studijní obor **Biomedicínské inženýrství a bioinformatika**

Ústav biomedicínského inženýrství

Student: Bc. Vít Plich

ID: 164993

Ročník: 2

Akademický rok: 2018/19

NÁZEV TÉMATU:

Detekce fibrilace síní v EKG

POKYNY PRO VYPRACOVÁNÍ:

1) Proveďte literární rešerši v oblasti automatické detekce fibrilace síní (FiS) v EKG záznamech. Popište parametry EKG a detekční metody, které lze pro tento účel použít. 2) Seznamte se s databází EKG signálů dostupnou na ÚBMI. Proveďte výpočet vybraných parametrů z EKG zdravých jedinců a pacientů s FiS. 3) Pomocí statistických testů určete parametry vhodné pro automatickou detekci FiS. 4) Navrhněte detektor FiS za použití vypočtených parametrů EKG a vybraného klasifikačního algoritmu. Navržený postup realizujte v prostředí Matlab. 5) Porovnejte výsledky s dostupnými údaji z literatury. Dosažené výsledky podrobně diskutujte.

DOPORUČENÁ LITERATURA:

[1] CURTIS, Michael J., Jules C. HANCOX, András FARKAS, et al. The Lambeth Conventions (II): Guidelines for the study of animal and human ventricular and supraventricular arrhythmias. Pharmacology and Therapeutics, Amsterdam: Elsevier Inc., 2013. roč. 139, vyd. 2, s. 213 – 248. ISSN 0163-7258.

[2] CLIFFORD, Gari D, Francisco AZUAJE a Patrick MCSHARRY. Advanced methods and tools for ECG data analysis. Boston: Artech House, 2006, 384 s. ISBN 1-58053-966-1.

[3] SAHOO, Sujit Kumar, Wenmiao LU, Sintiani Dewi TEDDY, Desok KIM, et al. Detection of atrial fibrillation from non-episodic ECG data: a review of methods. In 33rd Annual International Conference of the IEEE EMBS, Boston, USA, 2011, s. 4992-4995.

Termín zadání: 4.2.2019

Termín odevzdání: 17.5.2019

Vedoucí práce: Ing. Marina Ronzhina, Ph.D.

Konzultant:

prof. Ing. Ivo Provazník, Ph.D.
předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

Abstrakt

Tato diplomová práce se zabývá především detekci síňových pomocí HRV, klasifikaci Poincarého mapy a následnému rozdělení do dvou skupin pomocí neuronových sítí. Výsledkem je potom rozhodnutí, které veličiny jsou statisticky významné pro identifikaci síňových fibrilací, které ne, a klasifikace signálu na základě těchto veličin.

Klíčová slova

Detekce síňové fibrilace, AF, Poincare graf, neuronová síť, R-R entropie, SD 1, SD 2, ROC, AUC

Abstract

This diploma thesis deals with detection of atrial fibrillation from HRV, classification of Poincare map and in the end the divide into two groups, one with detected atrial fibrillation and one not. The result is the decision on which variables are statistically significant for the identification of atrial fibrillations and which are not, and classification of the ECG signals.

Keywords

Detection of atrial fibrillation, Fis, Poincare plot, HRV, SD1, SD2, R-R entropy, ROC, AUC

Bibliografická citace:

Plch, V. *Detekce fibrilace síní v EKG*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2019. 37 s. Vedoucí práce: Ing. Marina Ronzhina, Ph.D.

Prohlášení

„Prohlašuji, že svou závěrečnou práci na téma Detekce fibrilace síní v EKG jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené závěrečné práce dále prohlašuji, že v souvislosti s vytvořením této závěrečné práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne **17. května 2019**

.....

podpis autora

Poděkování

Děkuji vedoucímu diplomové práce Ing. Marina Ronzhina, Ph.D. za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé diplomové práce.

V Brně dne **17. května 2019**

.....
podpis autora

OBSAH

Obsah	7
úvod	1
1 Vznik elektrického napětí v srdci	2
1.1 Vznik membránového napětí	2
1.2 Převodní soustava srdce	3
1.3 Křivka EKG	4
1.4 Patologické změny v EKG	5
2 Detekce Atriální fibrilace pomocí intervalů R-R	7
2.1 Globální a lokální detekce pozice komplexů QRS	7
2.2 Úvod do analýzy dat	9
2.3 Analýza Poincarého mapy	10
2.4 Statistické vyhodnocení analyzovaných dat	14
2.5 Ověření přesnosti a funkce detektoru	23
2.6 Sestrojení detektoru	25
3 Závěr	29
Seznam symbolů, veličin a zkratk	34
Seznam příloh	35

Seznam obrázků

Obr. 1 Převodní systém srdce [2]

Obr. 2 Složení signálů EKG z jednotlivých signálů částí převodní soustavy

Obr. 3 Znázornění druhů signálů [12]

Obr. 4 Poincarého mapa pacienta bez síňové fibrilace (vlevo) a pacienta se síňovou fibrilací (vpravo)

Obr. 5 Poincarého mapa pacienta klasifikována jako bez síňové fibrilace (ektopický signál)

Obr. 6 Grafy SD 1: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

Obr. 7 Grafy SD 2: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

Obr. 8 Grafy SD 2: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

Obr. 9 Grafy Směrodatné odchylky RR: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

Obr. 10 Grafy střední hodnoty RR: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

Obr. 11 Grafy krokování (stepping): Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

Obr. 12 Grafy rozptylu (dispersion): Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

Obr. 13 Grafy VAI: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

Obr. 14 Grafy Entropie: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

Obr. 15 Závislost Směrodatné odchylky SD 2 [ms] na odchylce SD 1 [ms] detekované doktorem a pomocí funkce neuronových sítí v prostředí matlab (zelené – TP, modré – TN, červené – F)

Obr. 16 Závislost rozptylu [ms] na krokování [ms] detekované doktorem a pomocí funkce neuronových sítí v prostředí matlab (zelené – TP, modré – TN, červené – F)

Obr. 17 Závislost průměrného RR [ms] na směrodatné odchylce RR [ms] detekované doktorem a pomocí funkce neuronových sítí v prostředí matlab (zelené – TP, modré – TN, červené – F)

Obr. 18 Závislost Entropie [J/K] na VAI [°] detekované doktorem a pomocí funkce neuronových sítí v prostředí matlab (zelené – TP, modré – TN, červené – F)

Obr. 19 ROC trénovací skupiny dat (AUC = 0.980)

Obr. 20 ROC testovací skupiny dat (AUC = 0.957)

ÚVOD

Fibrilace síní je jednou z mnoha patologických disfunkcí srdečního svalu, která může mít za následek i smrt. Je tedy nezbytné fibrilaci síní odhalit co nejdříve, abychom mohli zamezit tragickým následkům této nemoci.

Tato práce se věnuje detekci fibrilací síní, respektive ověření několika známých principů této detekce a porovnáním výsledků těchto detekcí a implementací těchto principů do stromové rozhodovací struktury a struktury neuronových sítí v prostředí matlab. Výsledkem této práce je potom program, který rozhoduje, zda daný pacient trpí fibrilací síní nebo ne, rozhodnutí o tom, jakým způsobem je nejlepší detekovat fibrilace, popřípadě porovnání přesností a nepřesností jednotlivých metod nebo jejich kombinací získané nejpřesnější řešení. Práce se zabývá především řešením detekce pomocí intervalu R-R, znázorněním intervalů do Poincarého mapy a výpočtem z ní získaných veličin které slouží jako vstupy pro následující analýzy a klasifikace dat.

1 VZNIK ELEKTRICKÉHO NAPĚTÍ V SRDCI

Tato kapitola je čerpána především ze zdroje [1].

Na všech živých buňkách je možné změřit elektrické napětí. Jeho velikost se liší podle druhu buňky ale i podle jejího stavu. U srdečních buněk se velikost tohoto napětí v klidu pohybuje v rozmezí -70 až -80 mV (se záporným nábojem uvnitř buňky. Toto napětí vzniká následkem rozdílné propustnosti iontových kanálů v membráně buňky. A rozdílnou koncentrací vně a uvnitř buňky. Vzrušivé buňky, mezi které patří i buňky myokardu, Využívají změn tohoto napětí k šíření informací. Změna klidového napětí buňky nad její prahovou hodnotu vyvolá vznik akčního napětí. Tvar průběhu tohoto napětí potom určuje u jednotlivých buněk sílu kontrakce a dobu jejího trvání.

1.1 Vznik membránového napětí

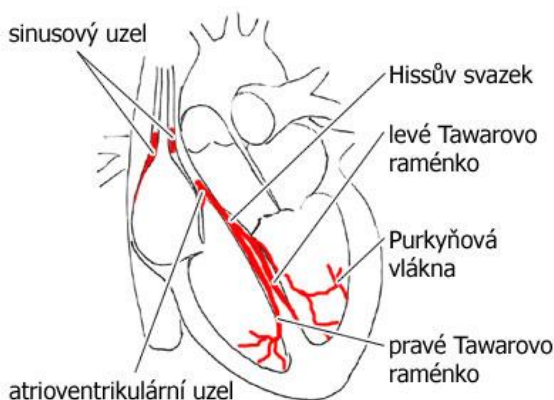
Membránové napětí vzniká na fosfolipidových membránách všech živých buněk. Toto napětí je způsobeno především rozdílnou koncentrací iontů vně a uvnitř buňky. Tuto rozdílnou koncentraci umožňují bílkovinné útvary prostupující skrz membránu. Nazýváme je iontovými kanály. Protože hlavním kationtem intracelulárního prostředí je draslík a hlavním aniontem extracelulárního prostředí je draslík, je nejdůležitější sodíko-draslíkový kanál, také nazývaný sodíko-draslíková pumpa. Pumpa jej nazýváme proto, že neustále pumpuje ven z buňky sodík, a naopak do buňky draslík, protože se jedná o pohyb proti koncentračnímu spádu, spotřebovává tato pumpa neustále energii dodávanou ve formě ATP (adenosintrifosfát). Dalším důležitým kanálem je sodíko-vápníkový výměnný mechanismus, při němž se využívá koncentračního spádu sodíku, který se pohybuje směrem dovnitř buňky, a oproti němu se dostává vápník z buňky ven. Tato výměna se používá především v době kontrakce srdečního svalu, kdy je v intracelulární tekutině velké množství vápníkových iontů. Mimo kontrakci je však vápníku v buňkách přítomno méně a nedochází k výměně tak často.

Klidové napětí je na buňkách udržováno pomocí difúze přes iontové kanály vždy ve směru koncentračního spádu. Největší podíl na tomto jevu mají ionty draslíku, protože draslíkový kanál se na membráně vyskytuje nejčastěji. Po nějakém čase dojde k ustálení a vyrovnání sil koncentračního spádu a elektrostatických sil, působících v opačném směru proti koncentračnímu spádu. Tato rovnováha nastává právě při napětí mezi -70 až -80 mV, při záporném náboji uvnitř buňky. U síňového myokardu se potom pohybuje v rozmezí od 65 do 70 mV. Pokud dojde k podráždění buňky vnějšími vlivy, dojde k šíření akčního potenciálu. U srdeční buňky k tomuto podráždění dochází elektrickým impulzem. Přesněji jde o depolarizaci o přibližně +5 mV. Následně se otevřou sodíkové kanály, jimiž začnou do buněk prudce proudit

ionty sodíku, což zvedne napětí na membráně na +20 mV. Když při této velmi rychlé reakci dosáhne napětí hodnoty přibližně -35 mV, dojde k otevření vápníkových kanálů. Tok vápníku je pomalý, a proto dokáže udržet napětí na membráně na hodnotě kolem 0 mV po dobu mezi 150 až 200 ms. Tuto dobu nazýváme fází plató a je typická pro srdeční buňky. Po této době se sodíkové i vápníkové kanály spontánně uzavírají a odeznívají i jím příslušné proudy. Aktivace a inaktivace iontových kanálů jsou tedy časově i napětově závislé děje.

1.2 Převodní soustava srdce

Schopnost vést akční potenciál mají všechny srdeční buňky i přes to, se některé z nich v průběhu vývoje přizpůsobily k vedení proudu na úkor kontraktibility. Tyto buňky tvoří tkáň, které nazýváme převodní soustava srdce. Do této soustavy patří postupně od místa vzniku signálu po směru jeho šíření (Obr. 1): sinoatriální uzel (SA), Hissův svazek, atrioventrikulární uzel (AV), levé a pravé Tawarovo raménko a Purkyňova vlákna.



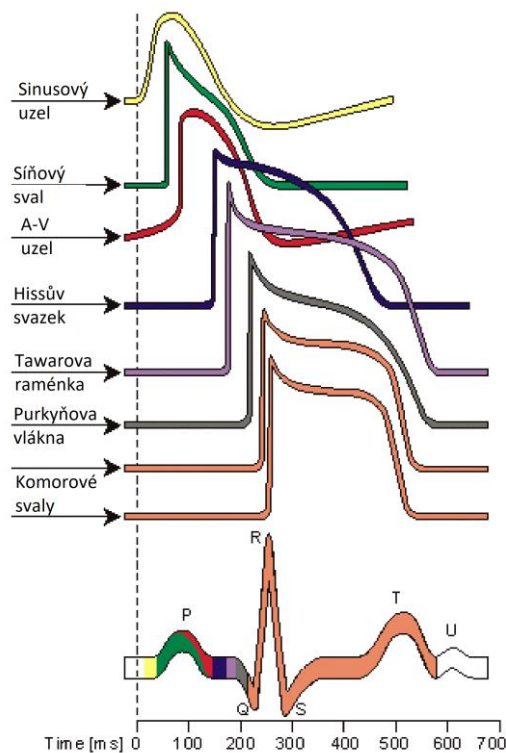
Obr. 1 Převodní systém srdce [2]

Za normálních podmínek je počáteční vzruch generován v SA uzlu a šíří se všemi směry přes stěny síně k AV uzlu. Pokud by došlo k disfunkci SA uzlu je jeho funkce nahrazena AV uzlem a Hissovým svazkem, směr šíření vlny v síních se ale obrátí, a klidová tepová frekvence se zpomalí. AV uzel, Hissův svazek, Tawarova raménka a Purkyňova vlákna tvoří dohromady komorovou část převodní soustavy. AV uzel, ležící pod endokardem, na spodině pravé síně nad trojcípou chlopní, přechází svou dolní částí volně v Hissův svazek. Hissův svazek prochází přes nevodivou přepážku až na mezikomorové septum. Společně se nazývá AV uzel a Hissův svazek jako síňokomorová (AV) junkce. V Hissově svazku také dochází ke zpomalení, rozdělení a izolaci drah pro jednotlivé komory.

1.3 Křivka EKG

Pokud sečteme akční potenciály generované jednotlivými buňkami převodního systému, a vezmeme-li v úvahu i jejich zpoždění v čase, bude výsledkem elektrokardiogram, tak jak je vidět na Obr. 2.

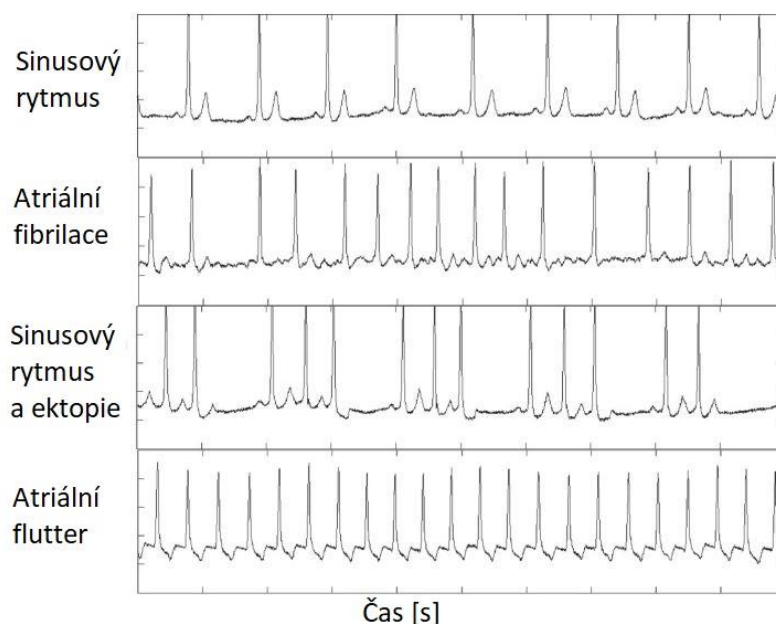
Z tohoto obrázku lze poznat, která část srdce způsobuje vlny a komplexy. Pokud jdeme zleva, počínaje vlnou P lze pozorovat že je ze začátku tvořena Sinusovým uzlem, z největší části ji tvoří signál síňového svalu a její konec tvoří A-V uzel. Změny na Hissově svazku nebo Tawarových raménkách se projeví na intervalu P-Q. Změny na Purkyňových vláknech se projeví na kmitu Q. Zbytek, tvořený z části vlnou Q a vlnami R, S a T, je tvořen z největší části komorovým svalem, a ostatní jevy v této době svou amplitudou mnohonásobně převyší.



Obr. 2 Složení signálů EKG z jednotlivých signálů částí převodní soustavy

1.4 Patologické změny v EKG

Vlna P je převážně kladná kulovitá vlna s výškou v rozmezí 0,1 až 0,15 (maximálně však 0,25) mV s trváním kratším než 100 ms. Pokud nalezneme vlnu P kladnou ve všech svodech jedná se s největší pravděpodobností o sinusový rytmus. Alternující vlna P s frekvencí 50-60 min⁻¹ jedná se o projev putujícího pacemakeru. U vyšších frekvencí se jedná o multifokální zrychlený sinusový rytmus. Pokud v některých svodech nalezneme vlnu P negativní jedná se nejspíše o junkční (nodální) rytmus (vycházející z AV junkce). Pokud se výskyt vlny neprokáže, jedná se nejčastěji právě o vadu způsobenou síňovou fibrilací. Muže se však jednat i o další poruchy jako je například flutter síní, sinoatriální blokádu, komorovou tachykardii, supra ventrikulární tachykardii, fibrilaci komor nebo flutter síní. Dalším patologickým projevem může být pozměnění intervalů mezi údery srdce. Tyto údery pak nenastávají tak pravidelně jako je tomu u sinusového EKG ale podobně jako u sinusového EKG s ektopií (Obr. 3). Tato podobnost může způsobit chyby klasifikace signálu z RR intervalů..



Obr. 3 Znáznornění druhů signálů [12]

Interval P-Q měří se od začátku vlny P a končí na začátku kmitu Q, nebo taky komplexu QRS. Tento úsek udává dobu nutnou k doputování signálu od sinoatriálního uzlu ke svalovině komor. Považujeme ho za izoelektrický, protože z něj určujeme nulovou linii signálu. Trvání tohoto intervalu je za normálních podmínek v rozmezí 120–200 ms. Kratší interval nasvědčuje tachykardii (zrychlené tepové frekvenci) a delší naopak o bradykardii (zpomalené tepové frekvenci)

Komplex QRS je zobrazením průběhu depolarizace komor. Tato depolarizace je postupná, jako první probíhá depolarizace septa z levého Tawarova raménka, pak

jsou depolarizovány subendokardiální vrstvy obou komor a vzruch se odtud šíří přes pracovní myokard až k epikardu. Komplex obsahuje obvykle 3 kmity P, Q a R. Normální trvání komplexu QRS je v rozmezí 60-110 ms. Prodloužení intervalu nad tyto hodnoty může být způsobeno bloádami Tawarových ramének, předčasnou komorovou excitací nebo idioventrikulárním rytmem.

Úsek ST je obvykle izoelektrický interval mezi koncem komplexu QRS a začátkem vlny T. Podstatou tohoto úseku je repolarizace komor (fáze plató). Počáteční bod tohoto úseku nazýváme jako junkční (značíme J) a normálně se tento bod vyskytuje ve stejné linii s počátkem komplexu QRS. V některých svodech je uváděna tolerance 0,05 mV. Elevace nebo deprese (zvýšení či snížení) úseku ST jsou považovány za projevy infarktu myokardu v akutním stádiu.

Vlna T je stejně jako úsek ST projevem repolarizace komor. Za normálních podmínek není vlna T kulovitého tvaru jako vlna P, ale má pozvolnější nástup a strmý sestup. Většinou je však také kladná, trvá přibližně 200 ms a její amplituda dosahuje 0,2-0,8 mV. Typická patologická vlna t pak může být například hluboká negativní vlna (vice než 0,5 mV), plochá nebo plošně negativní vlna, asymetrická negativní vlna nebo vysoká špičatá vlna T s úzkou základnou.

Vlna U je nalezená většinou 200-400 ms po ukončení vlny T. Její význam však není zcela jistý. Nález této vlny bývá fyziologický u sportovců a mladších pacientů.

2 DETEKCE ATRIÁLNÍ FIBRILACE POMOCÍ INTERVALŮ R-R

Základem této práce je rozlišit vstupní data s fibrilací síní od ostatních. Jednotlivé metody jsou testovány na vstupních datech dostupných od firmy BTL Zdravotnická Technika, a.s. Těchto signálů bylo původně 6884 avšak některé jsou nevhodně nasnímané, překryté rušením nebo jinými chybami. Po vymazání nevhodných signálů zbylo 6830 použitelných signálů, z toho je 442 signálů určených odborným lékařem jako postižených síňovou fibrilací (AF) a 6238 jako signálů bez AF. Signály bez AF však můžou být zatíženy jinými srdečními vadami. Data jsou snímány v klinickém prostředí s vzorkovací frekvencí 500 Hz a délkou signálu 10 s. Tyto data byly v první části převedeny pomocí skriptu *AFIB_BTL_dataLoading.m* z formátu XML do formátu matice a uloženy jako *18570_to_25455_non_AF_UBMI.mat*, která obsahovala pouze EKG signály s nedetekovanou síňovou fibrilací, a matice *243_detected_AF_UBMI.mat*, která obsahovala pouze data s klasifikovanou a lékařem potvrzenou síňovou fibrilací. V druhé části jsou ale data přepracována, jsou přesněji určeny kategorie signálů (u některých chyběla klasifikace a tak byly mylně pokládány za signály bez AF), a to tak, že jsou z dat vyjmuty pouze pozice R vln detekované přímo elektrokardiografem firmy BTL uložených pomocí funkce *grsSave.m* do souboru *BTL_qrs.mat* a je použita tabulka *indAfibTrainTest.mat* ve které jsou uloženy finální klasifikace pro jednotlivé signály a jejich zařazení do trénovací nebo testovací skupiny. Při Praktické části jsou potom signály roztříděny do těchto skupin a náhodně promíchány tak, aby simulovali náhodné uskupení návštěvníků kliniky.

Pro detekci pomocí R-R intervalů je vždy nutné nejdříve detekovat pozice vln R. Zde je využito pozic přímo detekovaných elektrokardiografem, avšak jinak se využívá detekce pomocí globální či lokální funkce využívající buď pouze jeden svod, nebo všechny dostupné svody. Je však ještě možnost využít jednoho svodu vytvořeného reprezentací ostatních signálů. Může ale nastat problém, že v jednotlivých svodech se vlna R promítne vždy v trochu odlišných časových momentech, a to kvůli rozdílnému umístění elektrod, ze kterých vyplývá rozdílná vzdálenost od daného místa a rozdílná doba šíření signálu k elektrodě. Proto této metody při této detekci nevyužíváme.

2.1 Globální a lokální detekce pozice komplexů QRS

Globální i lokální detekce byla v první části řešena pomocí vlnkové transformace signálu. A to pomocí funkce *QRSdetektor_global.m* a *QRSdetektor_local.m* vydané Ing. Martinem Vítkem roku 2011. Globální funkce byla řešena pouze úpravou

lokální. V druhé části je potom použita detekce R vln přímo pomocí vnitřní funkce elektrokardiografu.

Dekompozice pomocí kontinuální vlnkové transformace je operace, při které dochází k rozkladu signálu na jednotlivá pásma. A to tak, že předpokládáme vlnkovou transformaci ve spojitém čase. Díky níž získáme časově-frekvenční popis signálu. Tím získáme informaci o frekvencích signálu a jejich výskytu. Této transformace se využívá právě pro nalezení míst výskytu, což je výhodou oproti Fourierově transformaci. Tuto vlnkovou transformaci lze definovat vztahem [4]:

$$y(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t-b}{a} \right) dt \quad a, b \in \mathbb{R}; a \neq 0, \quad 2.1$$

Z uvedeného vztahu je patrné, že se jedná o časově-frekvenční rozklad, který může být interpretován jako korelace originálního signálu $x(t)$ s vlnkami odvozenými z vlnky mateřské $\psi(t)$. Tato mateřská vlnka musí mít nulovou střední hodnotou a také musí být obecně komplexní. Název vlnky vyplývá z tvaru funkce $\psi(t)$. Symbol $*$ označuje komplexně sdruženou funkci, protože vlnky obecně mohou být komplexní. Výslednou funkci $y(a, b)$ a také vlnky $\psi_{a,b}(t)$ lze popsat dvěma spojitě proměnnými parametry: parametr a ovlivňuje časovou dilataci a parametr b slouží k zajištění posunu v čase. Při správném pochopení problematiky můžeme tedy říci, že je-li konstanta a větší, básová funkce je širší a tím nám přináší informaci o nízkofrekvenčních komponentách signálu a naopak. Je tedy patrné, že čím je vyšší frekvence, tím je vyšší i časové rozlišení. Konstanta \sqrt{a} normalizuje energii jednotlivých vlnek.

Lokální detekce je řešena pomocí vlnky *Bior 1.5* předdefinované v prostředí matlab. Při dekompozici vstupního signálu pomocí této vlnky je vybráno předdefinované pásmo zvolené jako nejvhodnější k detekci z lidského EKG, dále jsou nalezeny extrémy a mezi nimi vždy průchody nulou. Tyto průchody nulou jsou po eliminaci zpoždění detekce považovány za lokální vlny R. U globálního detektoru potom dochází k tomu, že z jednotlivých svodů se detekují pozice vln R a pomocí mediánu je vybrána hodnota nejlépe odpovídající skutečné časové pozici vlny R.

V poslední části klasifikace signálu jsme však využily signál R vln, generovaný přímo snímacím přístrojem EKG, a to pro lepší aplikaci v budoucnosti přímo pro daný přístroj. Vadou těchto signálů je však jejich délka, která je pro některé výpočty až příliš krátká. To jsme řešili periodickým opakováním signálu u konkrétních veličin. Tím by nemělo dojít k narušení hodnot počítaných ze signálu a zároveň se zlepší výpočet některých hodnot jako je například entropie, která je počítána pomocí posuvného okna o určité délce m , která vyžaduje určitou minimální délku signálu a minimální počet R-R intervalů. Bylo vyzkoušeno několik variant, a to jak

ponechání původní délky signálu, tak její N -násoben. Nejlepší výsledky pak byly změřeny právě pro $N=2$.

2.2 Úvod do analýzy dat

Analýza dat je jednou z nejdůležitějších částí klasifikace signálu. Jde v ní především o studování souvislostí dat. Nejvšeobecnější název takovéto analýzy se dá shrnout jako „rozpoznávání obrazů“ přeložením z anglického výrazu „pattern recognition“. Jedná se však spíše o volný překlad, jelikož doslovný překlad by byl „rozpoznávání vzorů (schémat, šablon)“. Ty můžou být hodnoceny pomocí mnoha způsobů, například grafů, vektorů hodnot či soubory různých veličin. Jiným názvem mohou být například „vícerozměrné statistické metody“. Jde vlastně o soubor získávání různých spojitostí, či skrytých jevů, mezi určitými soubory dat a následným rozřazením do skupin pomocí různých spojitostí v rámci jednotlivých skupin. Toto získávání veličin, které pomáhají rozdělit jednotlivé skupiny dat můžeme nazývat jako „data mining“ což je v překladu dolování či vytěžování dat. Jiný přístup je potom „machine learning“ rozšířený hlavně v poslední době. V překladu se tedy jedná o strojové učení. Což znamená že část práce s nalezením podobností se dá implementovat do počítačových programů. Které dokážou samy vytěžit část dat. Další možností jsou také neuronové sítě. Které jsou přesněji rozebrány až v dalších kapitolách této práce.

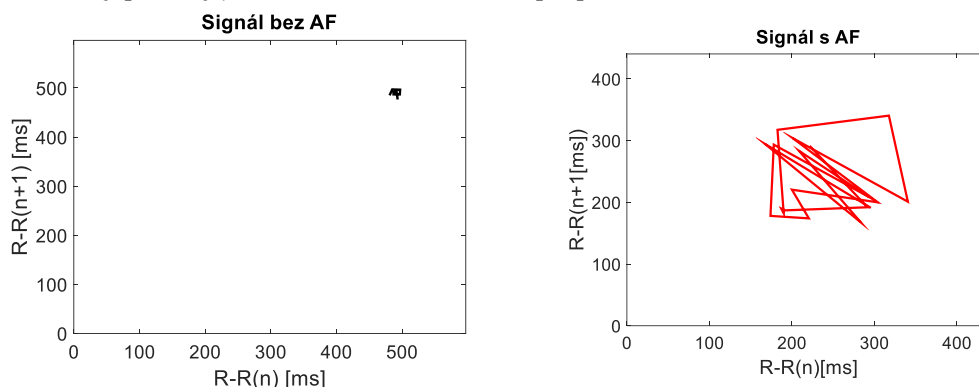
Cílem Analýzy dat je tedy převážně posouzení zkoumaného objektu, na základě kterého je možné buď rozhodnout o typu či charakteru daného objektu, posoudit kvalitu tohoto objektu, anebo rozhodnutí o budoucí akci. Teoreticky se však jedná o stejnou nebo alespoň velmi podobnou úlohu. Data, která jsou na začátku analýzy, se dělí na dvě skupiny. A to data buď spojitá (váha, délka a další) která je potřeba většinou před analýzou diskretizovat (kvantování, vzorkování), či data přímo diskrétní (barva, rasa, počet okvětních lístků, a podobné).

Dále je nutné nalézt vztahy mezi veličinami, to ale znamená nalezení zákonitostí a podstat v rozloženích hodnot, závislostí mezi hodnotami použitých proměnných nejlépe i nalezení matematického vztahu. Výsledky analytických výpočtů potom slouží k předzpracování dat ke klasifikaci, ale mohou být i přímo výsledkem zpracování, to například pokud výpočet přímo udává kvalitu předmětu. Obecně však způsob, jak určit veličiny nesoucí nejvíce informace pro klasifikaci, není teoreticky formalizován [15][16]. Takže neexistuje teorie, podle které je možné předem určit, které veličiny nesou důležitou informaci a které veličiny ne. Dnešní teorie však mohou poskytnout částečné řešení tohoto problému. Mohou pomoci buď vybrat minimální počet pro určitou přesnost ze souboru, nebo mohou pomoci vyjádřit původní skupinu veličin pomocí menšího počtu skrytých (latentních) proměnných, které nejde přímo změřit, avšak mohou být věcně interpretovány.

Konečným výsledkem analýzy potom bývá klasifikace, klasifikací rozumíme rozdělení jisté vstupní množiny objektů do konečného počtu skupin, který má dostatečně podobné společné vlastnosti (vstupní proměnné). Pro to je však nutné předem stanovit jistá klasifikační kritéria, například předmět musí patřit pouze do jedné skupiny nikoliv do více skupin naráz. Klasifikace jsou prováděny vždy pomocí takzvaného klasifikátoru. Klasifikátor je postup (algoritmus), na vstupu kterého jsou data, popisující analyzované objekty, a jeho výstupem je potom zařazení do jednotlivých skupin na základě těchto vstupních dat.

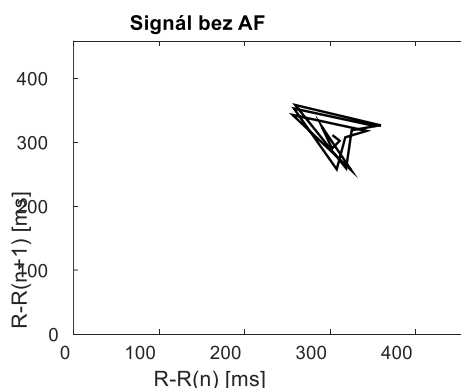
2.3 Analýza Poincarého mapy

Analýza Poincarého mapy je geometrická nelineární analýza sloužící k popisu dat EKG z detekovaných vln R. Jedná se o grafickou reprezentaci časové posloupnosti vždy tří po sobě jdoucích úderů srdce do grafu a to tak, že dva po sobě jdoucí intervaly mezi třemi vlnami R jsou vyneseny do grafu jako jeden bod. Na ose x má vždy hodnotu délky intervalu R-R (n) a na ose y potom následujícího intervalu R-R (n+1). [7] Na Obr. 4 (vlevo) je vidět Poincarého mapa zdravého pacienta a na Obr. 4 (vpravo) je vidět Poincarého mapa při síňové fibrilaci.



Obr. 4 Poincarého mapa pacienta bez síňové fibrilace (vlevo) a pacienta se síňovou fibrilací (vpravo)

Z principu této mapy je zřejmé, že čím delší úsek EKG máme k dispozici, tím jsou mapy vypovídající. Pokud máme velmi krátký úsek, může být graf nevypovídající.



Obr. 5 Poincarého mapa pacienta klasifikována jako bez síňové fibrilace (ektopický signál)

Z dat dostupných v této práci je potom zřejmé, že ne vždy se v tomto grafu musí

fibrilace s ní projevít stejnou změnou grafu. Například na Obr. 5 vidíme Poincarého mapu, která ukazuje signál, na němž nebyla klasifikována síňová fibrilace, avšak takovouto mapu připomíná. Jednou z možností zmiňovanou v článku [6] je detekce tohoto typického útvaru. Hlavním problémem špatné klasifikace v tomto případě však nejsou jen problémy s podobností map, ale především délka signálu. V rozsahu 10 sekund, jsme schopní nasnímat mezi 7 až 30 R vlnami (při tepové frekvenci předpokládané mezi 42 až 180 tepy za minutu). Lépe se proto určují delší signály, například trvající 5 minut, jako v literatuře [6]. Některé snímané signály mají však I méně než 4 R vlny. Proto jsme 16 signálů, které minimální délkou nevyhovují, vyřadili z klasifikace. Může jít jak o špatný výpočet vlny R, tak o signál EKG znehodnocený šumem. Délku ostatních signálů jsme potom při některých výpočtech museli zdvojnásobit, a to tak že vypočtené intervaly R-R jsme kopírovali dvakrát za sebe. Tím sice nepřibude signálu žádná nová informace, avšak se bavíme neurčitých hodnot při výpočtu s posuvným oknem, které požaduje pro svou funkci minimální délku signálu. Tím je docíleno menší chybovosti než dopočítáváním náhodných podobných hodnot intervalů R-R.

Pro jednoduchost se snažíme popsat graf pomocí jeho číselných reprezentací. Nejčastěji se uvádí směrodatné odchylky ve směru pod úhlem 45°, označená jako SD_1 podle vzorce 2.2 [5], a 135°, označená jako SD_2 podle vzorce 2.3 [5]. Dále je vypočítaná plocha oválu S tvořená odchylkami SD_1 a SD_2 podle vzorce 2.4 [5]. Dále průměrný nárůst intervalu R-R na krok (stepping) podle vzorce 2.5 [6] a rozptyl kolem středového bodu (dispersion) podle vzorce 2.6 [6].

$$SD_1 = \sqrt{\text{std} \left(\frac{RR_i - RR_{i+1}}{\sqrt{2}} \right)}, \quad 2.2$$

$$SD_2 = \sqrt{\text{std} \left(\frac{RR_i + RR_{i+1}}{\sqrt{2}} \right)}, \quad 2.3$$

$$S = \pi \cdot SD_1 \cdot SD_2 \quad 2.4$$

Kde $\text{std}(x)$ je směrodatná odchylka (standard deviation), RR_i je vektor R-R intervalů 1 až $n-1$ a RR_{i+1} je vektor R-R intervalů 2 až $n+1$.

$$\text{stepping} = \frac{\frac{1}{n-2} \sum_{j=1}^{n-2} \sqrt{(I_j - I_{j+1})^2 + (I_{j+1} - I_{j+2})^2}}{\frac{1}{n} \sum_{j=1}^n I_j}, \quad 2.5$$

$$\text{dispersion} = \sqrt{\frac{\frac{1}{2(n-1)} \sum_{j=1}^{n-1} (I_j - I_{j+1})^2 - \left(\frac{1}{(n-1)\sqrt{2}} \sum_{j=1}^{n-1} |I_j - I_{j+1}| \right)^2}{\frac{1}{2(n-1)} (-I_1 - I_n + 2 \cdot \sum_{j=1}^{n-1} I_j)}} \quad 2.6$$

Kde n je počet bodu na Poincarého mapě, I_j je pozice j -tého bodu a I_{j+1} je pozice následujícího bodu.

Další získanou veličinou je VAI (vector angular index) který je podobný, ne však stejný jako SD_2 . Jedná se o průměrnou odchylku ve stupních od úhlu 45° pro daný signál. A je vypočítán podle vzorce 2.7 [8]. Tato veličina reprezentuje odchylky úhlů jednotlivých bodů, svíraných mezi spojnicí bodu v Poincarého plotu s počátkem a osou x , od úhlu 45° . Z těchto hodnot je potom vypočítán aritmetický průměr. Hodnota je založena na předpokladu, že body v Poincarého plotu jsou umístěny blíže k úhlu 45° jako je znázorněno například na Obr. 4.

$$VAI = \frac{\sum_{i=1}^N |\theta_i - 45|}{N} \quad 2.7$$

Nejdůležitější veličinou je potom entropie signálu. Tato veličina sama o sobě zaručuje efektivní a poměrně přesnou klasifikaci atriální fibrilace [14]. Teorie entropie vychází ze základních zákonů fyziky. První zákon je zákon o zachování energie, který říká že v uzavřeném systému energie nevzniká ani nezaniká ale pouze mění svůj stav. Druhý zákon termodynamiky říká, že se mechanická práce přeměňuje na teplo ale ne opačně. Při všech změnách energie se část energie systému mění na teplo a dochází k postupné ztrátě energie a tím k vzrůstu entropie. Z tohoto vztahu je entropie definována jako extenzivní složka tepla a popisuje degradaci tepla nastávající u nezvratných změn. Tuto degradaci lze chápat jako postupně ubývajících schopnost práce. Entropie může v uzavřeném systému tedy pouze zůstat beze změny nebo stoupat a je též stavovou veličinou systému. [9]

Je však více možností, jak chápat entropii, a to třeba podle teorie informace. Kde je entropie jako míra neurčitosti (neuspořádanosti) systému [10]. Pokud budeme brát graf HRV (tachogram) jako zprávu, která nese informaci u stavu systému (buď má pacient atriální fibrilaci nebo ne) je vidět v signálu HRV určitá uspořádanost u pacientů s AF (pravidelnost po sobě jdoucích vzdáleností intervalů R-R, postupné narůstání, klesání) a u pacientů bez AF je vidět jistá nahodilost (neuspořádanost, nepravidelnost, nahodilost) signálu HRV. Nejběžnějším využitím entropie je její odhad [11]. Tuto metodu pro výpočet přibližné entropie (AppEN) navrhnul Pincus na krátkých zašuměných datech. Tato metoda přiřazuje nezáporná čísla pro signály s větší nepravidelností dat. Výpočet entropie určuje míru závislosti vstupu o délce N na předchozích m výstupech s tolerancí r , která určuje toleranci podobnosti dvou signálů. Tato metoda je v prostředí matlab interpretována funkcí *approxEnt* kde je nejdříve podle vzorce 2.8 vypočítána zpožděná rekonstrukce Y_{1-n} pro N vzorku dimenze m a potom podle vzorce 2.9 a 2.10 vypočítána přibližná entropie [21].

$$N_i = \sum_{i=1, i \neq k}^N 1(\|Y_i - Y_k\|_\infty < R) \quad 2.8$$

$$\Phi_m = (n - m + 1)^{-1} \sum_{i=1}^{N-m+1} \log(N_i) \quad 2.9$$

$$aproxEnt = \Phi_m - \Phi_{m+1} \quad 2.10$$

Vylepšením této metody je potom metoda vzorkové entropie (SampEn) [12] podle vzorce 2.11 [13]. Hlavními výhodami je že vzorková entropie nepočítá shody vzorků se sebou samými a to, že se nepoužívá při odhadu podmíněné pravděpodobnosti vzorec. Výsledky odhadu entropie a vzorkové entropie jsou velmi podobné.

$$SampEn(m, r, N) = -\ln \left[\frac{B^{m+1}(r)}{B^m(r)} \right] \quad 2.11$$

Další možností výpočtu entropie je víceúrovňová entropie (MSE) ve které je signál rozdělen pomocí koeficientů do bloků. Hodnoty entropie se pak počítají v úrovních signálu a jsou studovány v každé úrovni jednotlivě. Tato metoda je podobná zpracování signálu ve frekvenční oblasti. V této metodě je však potřeba podvzorkování signálu, a proto není vhodná pro velmi krátké signály, jako jsou v této práci, proto bylo použití této metody zamítnuto.

Pokud srovnáme tyto metody je přibližná entropie velmi závislá na délce signálu, což při velkých rozdílech v délkách signálů, jako jsou v datech od firmy BTL (až několikanásobné rozdíly), zavádí poměrně velkou chybu. Vzorková entropie je však navržena tak, aby právě tuto vadu přibližné entropie eliminovala. Zajímavé však je, že ačkoliv je každá následující entropie odvozena od předchozí udávají na sobě nezávislé výsledky. Nejdůležitější pro efektivitu entropie jako veličiny použité pro rozhodování při klasifikaci signálů je však volba parametru hranice podobnosti dvou vzorků – r . Proto je v pozdější části hodnota měněna v cyklu a jsou pozorovány výsledky klasifikace jako je plocha pod křivkou a F1measure.

Pro přesnější výpočet pomocí neuronových sítí je použita i průměrná délka intervalu R-R a směrodatná odchylka R-R intervalů použitých dat. Protože data obsahují velké množství EKG bez AF a pouze málo pacientů trpících AF je jako trénovací skupina pro neuronovou síť vybrána kombinace 150 pacientů s fibrilací síní a 1000 pacientů považovaných za zdravé, ve skutečnosti však všichni být zdraví nemusí, pouze nemají fibrilaci síní nebo jim nebyla klasifikována.

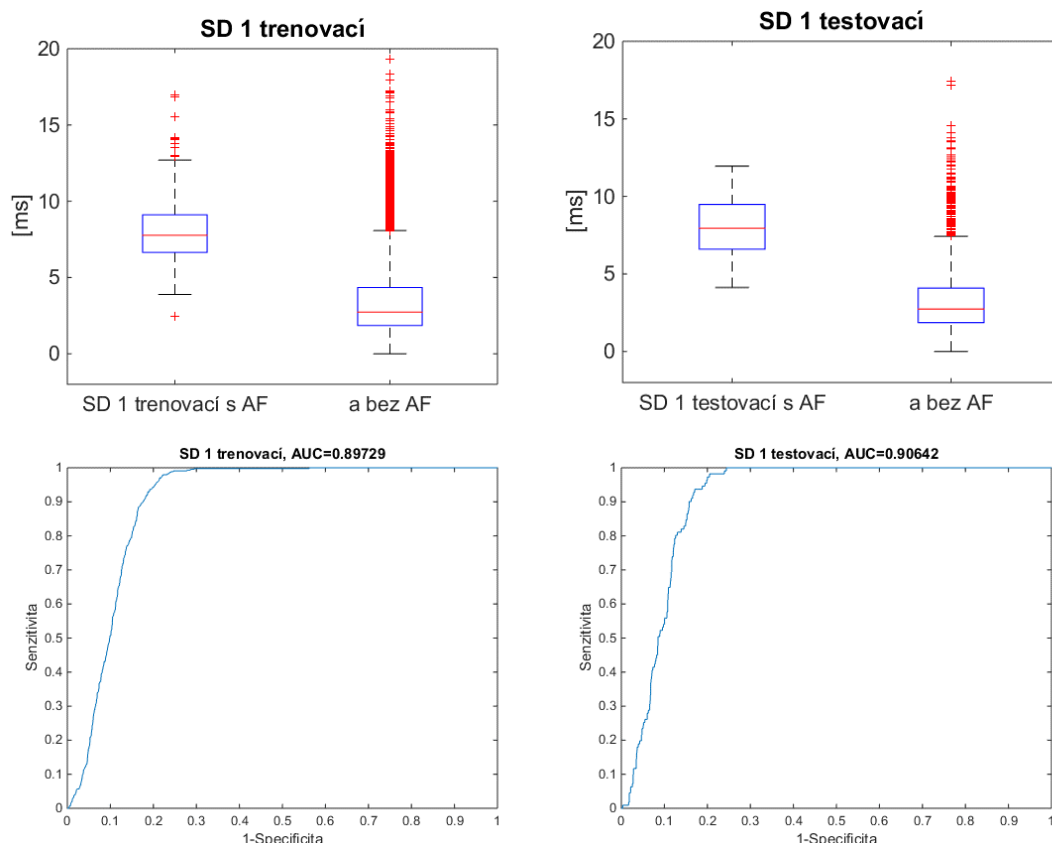
2.4 Statistické vyhodnocení analyzovaných dat

Abychom zjistily statistickou důležitost každé z numerických reprezentací Poincarého plotu, musíme provést statistické testování a testování rozložení dat. Prvním testem je vykreslení dat pomocí krabicového diagramu (box plotu) vždy pro každý parametr zvlášť a to tak, že vlevo se nachází data s detekovanou atriální fibrilací a vpravo bez detekované fibrilace. Tyto grafy vykreslují všechna dostupná data.

Krabicový diagram (vykreslený pomocí funkce *boxplot.m* a upravený pomocí funkce *boxploty.m*), jako je například na Obr. 6 nahoře vlevo, je schéma, které poskytuje v jednom grafu informaci o maximální a minimální hodnotě, mediánu a horním a dolním kvartilu vstupních dat. Dolní a horní strana obdélníku (krabice) reprezentují dolní a horní kvartil. Dolní kvartil je hodnota, pod kterou leží přesně čtvrtina všech hodnot, zatímco horní kvartil je hodnota, nad kterou leží přesně čtvrtina všech hodnot. Mezi těmito kvartily potom leží mezikvartilové rozpětí, v němž leží medián. Medián, na obrázcích vodorovná červená čára, dělí potom data přesně na dvě poloviny. Tyto tři úsečky potom dělí data na zhruba stejně čtyři části. Dolní svíslá úsečka (fous) potom odpovídá hodnotám ležícím pod krabicí, avšak pouze do maximální vzdálenosti jeden a půl násobku výšky krabice. Konec fousu (černá vodorovná čára) potom reprezentuje nejmenší hodnotu ze souboru. Naopak je to pak u horního fousu, který sahá také do jeden a půl násobku výšky grafu, a jeho horní vodorovná čára znázorňuje největší hodnotu souboru dat. Jak jde vidět z grafu jsou zde ještě červené body, ty reprezentují odlehlé hodnoty v grafu. [17]

Pro všechny data je potom vyhodnocen statistický nepárový t-test a kvůli jeho ovlivnitelnosti poměrem velikostí vstupních dat je vypočítán i Wilcoxonův test. Nepárový t-test je test pro data tvořené dvěma různými skupinami dat, která jsou vzájemně nezávislá. Což tomuto výběru dat vyhovuje. [18]

Pomocí Wilcoxonova testu (v matlabu *ranksum*) testujeme dva výběry, přičemž testujeme veličiny jejichž rozdělení nemusí odpovídat Gaussovu normálnímu rozdělení. Testujeme potom hypotézu rovnosti jejich distribučních funkcí, na základě ověřeného symetrického rozložení. Nulová hypotéza předpokládá shodné mediány obou těchto rozložení, a proto se ji snažíme zamítnout. Zamítnutí nulové hypotézy potom znamená, že mediány obou rozložení nejsou shodné. Pro zamítnutí nulové hypotézy tohoto testu je potřeba aby výsledek vypočítaný pomocí funkce byl roven 1.

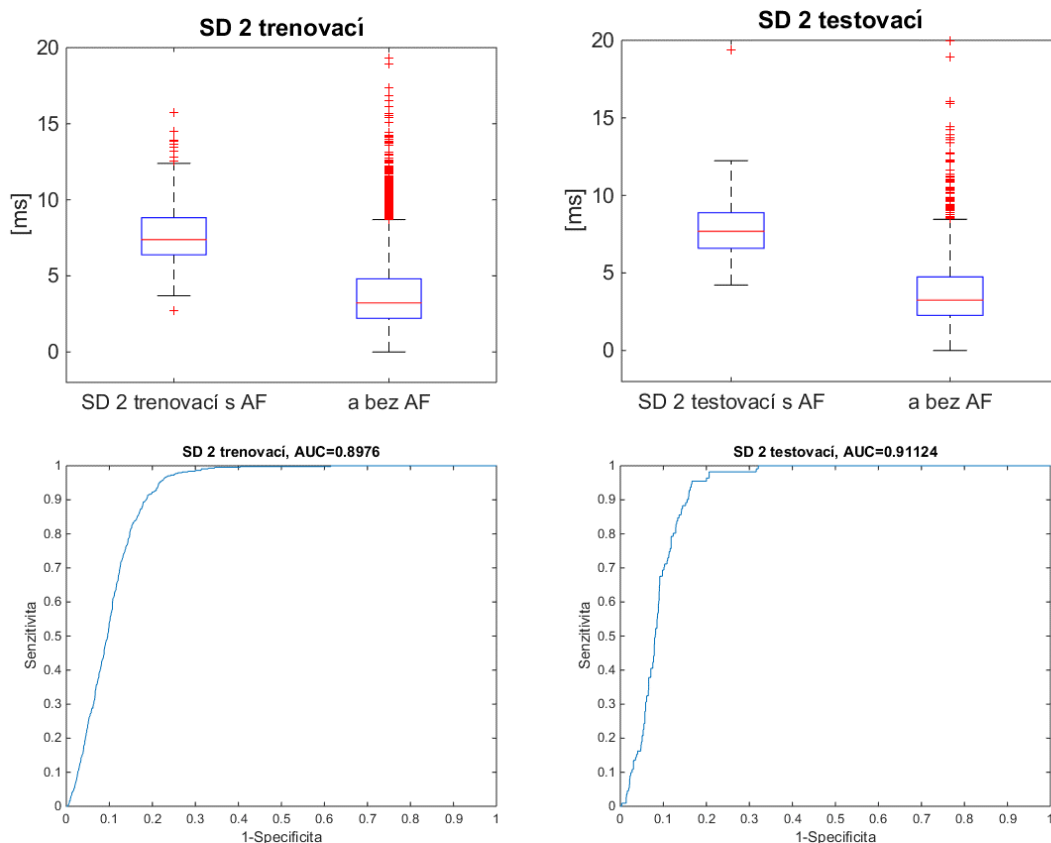


Obr. 6 Grafy SD 1: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

Jak lze vidět z box plotů na (Obr. 6) trénovací i testovací skupiny se překrývají. Největším problémem je zde velké množství odlehlých hodnot ve skupině bez detekovaných AF. Proto jsou trénovací data testována pomocí dvou-výběrového nepárového t-testu a wilcoxonova testu. A to vždy data s Fibrilací proti datům bez fibrilace pro každou reprezentaci zvlášť. Pro data jsme použily nulovou hypotézu, která předpokládá shodné střední hodnoty obou skupin. Pokud se nám podaří zamítnout nulovou hypotézu, na hladině významnosti $\alpha=5\%$, a přijmout hypotézu alternativní. Můžeme tedy prohlásit, že tyto dva výběry mají statisticky prokazatelně rozdílné střední hodnoty (i mediány). To znamená že se hodí k rozlišování dat zatížených AF od dat bez AF. V prostředí matlab k tomuto slouží funkce `ttest2`. Výsledek t-testu je $h=0$ a výsledek wilcoxonova testu je $h=1$. Na základě těchto testů jsme pro směrodatnou odchylku ve směru diagonály (SD 1) nulovou hypotézu zamítnuly. To znamená že Jsme přijmuly hypotézu alternativní, a můžeme prohlásit, že je SD 1 veličinou statisticky významnou pro rozlišování AF. Ve spodní části Obr. 6 Lze potom vidět, že AUC pro hodnotu SD 1 je přibližně 0.9, to znamená že hodnota sama o sobě poměrně dobře rozděluje obě tyto skupiny.

Další veličinou je směrodatná odchylka kolmá na diagonálu (SD 2). Box ploty pro tuto veličinu jsou zobrazeny na Obr. 7. Již na první pohled lze vidět jistotu

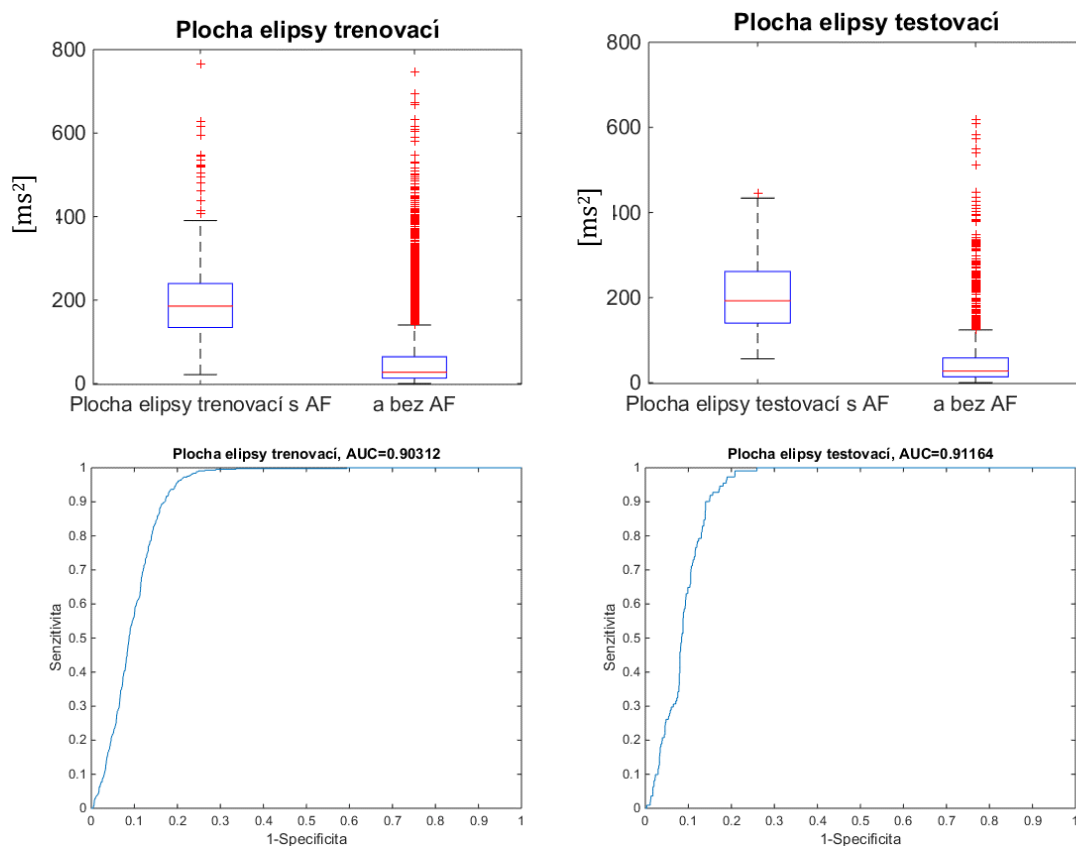
podobnost s odchylkou SD 1 až na malé rozdíly. U SD 2 lze pozorovat větší rozptýly u obou veličin, vzniká u nich i větší překryv. Po testování t-testem a wilcoxonovým testem jsme i pro směrodatnou odchylku SD 2 nulové hypotézy zamítly. To znamená že i směrodatná odchylka ve směru kolmém na diagonálu je statisticky významná na hladině významnosti $\alpha=5\%$ a i na základě této směrodatné odchylky můžeme rozhodovat, zda se u dat jedná o AF nebo ne. Podle ROC grafů ve spodní části Obr. 7 jde pak vidět že rozlišovací schopnost SD 2 je podobná jako SD 1.



Obr. 7 Grafy SD 2: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

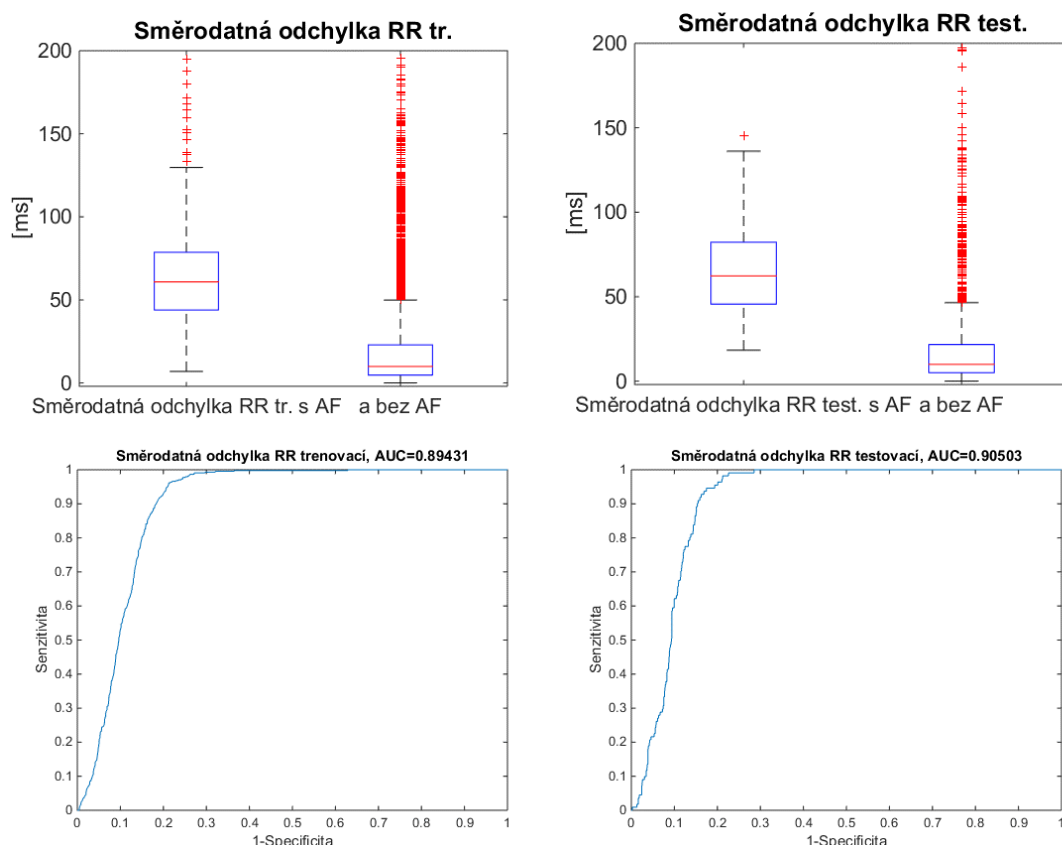
Z obou těchto odchylek se dá vypočítat plocha elipsy podle vzorce 2.4. Znovu jako v předchozích případech jsou zobrazeny box ploty na Obr. 8. Tyto Box ploty musely být upraveny, a to tak, že na ose y musely být pro viditelnost oříznuty na maximální hodnotu 800 $[ms^2]$ (dosahoval až 6000 pro odlehlé hodnoty). Jak je však z těchto grafů vidět, dosahují skupiny již většího překryvu mezi daty. To se promítlo i na výsledné pravděpodobnosti t-testu a wilcoxonova testu, které byly o několik řádů vyšší než u SD 1 a SD 2, avšak stále byla nižší než hraničních 5 %. V důsledku toho jsme tedy znovu zamítly nulovou hypotézy a přijmuly jsme alternativní. Takže i u plochy elipsy tvořené pomocí směrodatných odchylek SD 1 a SD 2 lze

předpokládat statisticky významnou odlišnost středních hodnot výběrů s AF a bez AF. Z grafů ROC jde potom vidět, že hodnota AUC je mírně vyšší než u SD 1 a SD 2.



Obr. 8 Grafy SD 2: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

O něco hůře pak vychází Směrodatná odchylka RR intervalů (Obr. 9). Tato veličina se nepočítá z Poincarého plotu, ale vychází čistě z hodnot RR intervalů signálu. Na Obr. 9 potom můžeme vidět znovu dva box ploty pro tuto veličinu.

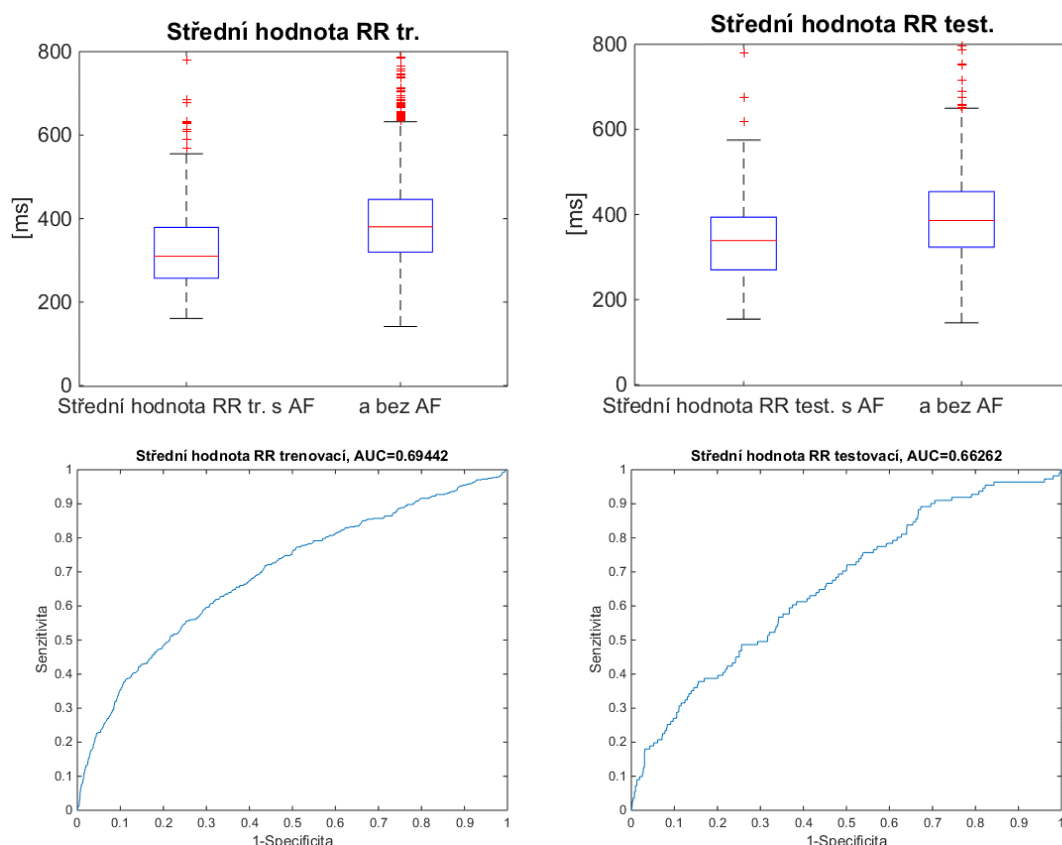


Obr. 9 Grafy Směrodatné odchylky RR: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

Tyto boxploty byly znovu oříznuty z důvodu extrémních odlehlých hodnot, dosahujícím až k 1500 [ms]. Jak můžeme z grafů vyčíst, tyto skupiny se pro směrodatnou odchylku intervalů R-R do velké míry překrývají, avšak mediány obou výběrů leží daleko od sebe. Výsledek t-testu i wilcoxonova testu nasvědčuje tomu, že ani tyto dvě rozložení nemají podobnou střední hodnotu (medián). Takže znovu zamítáme nulovou hypotézu.

Další testovanou veličinou je střední hodnota RR intervalů. Jak můžeme vidět na box plotech na Obr. 10 střední hodnota intervalu RR není pro detekci AF příliš vhodná. Již z těchto grafů lze vyčíst že se hodnoty velmi překrývají, a téměř celá množina dat s AF spadá do rozmezí od minima po maximum hodnot bez AF. Po vypočítání t-testu a wilcoxonova testu však znovu nulovou hypotézu zamítáme. Takže i tento parametr je vhodný k další analýze i když z parametrů vhodných k rozlišení signálů vychází nejhůře. Z hodnot AUC potom můžeme vyčíst, že se blíží náhodnému rozdělení. Avšak když jsme se pokusily tuto hodnotu z klasifikace vynechat klesla hodnota F_1 measure až o 0.02. I když to není příliš velký pokles, zdá se že každá hodnota, která není náhodná a přináší do metody novou informaci, má vliv na zvýšení hodnoty přesnosti klasifikace. Toto tvrzení podkládá i literatura [20] v níž je řešena klasifikace signálů pomocí 75 hodnot spojených v jednu komplexní

klasifikační metodu. Díky tomuto postupu byla celková přesnost podle hodnoty F_1 measure rovna 0.868. Je zde však také řešeno, že ani tato metoda měření kvality klasifikace nemusí být nejlepší. A tohoto výsledku bylo docíleno právě ve zpětné vazbě na hodnotu F_1 measure.

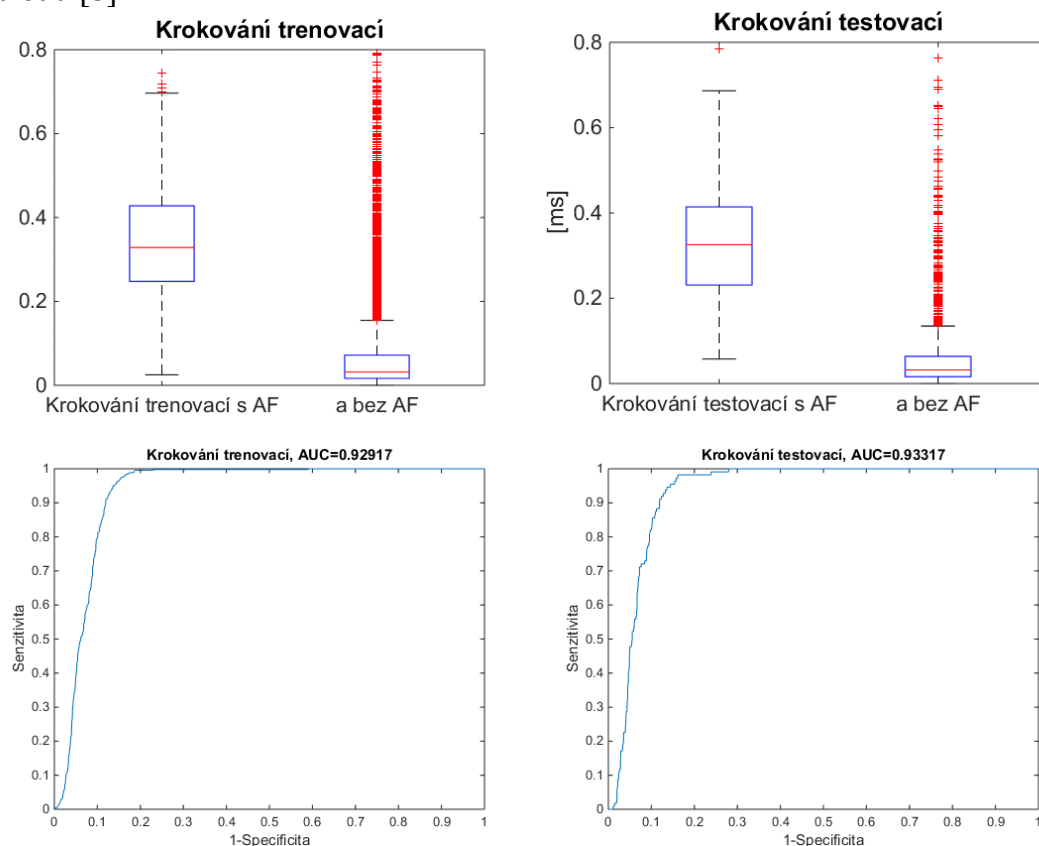


Obr. 10 Grafy střední hodnoty RR: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

Nejlépe ze všech potom dopadlo krokování. Tato hodnota udává průměrný nárůst vzdáleností dvou bodů a je dopočítána podle vzorce 2.5 [6]. Její box ploty jsou zobrazeny na Obr. 11. Tento graf musel být znovu oříznut. Lze na něm vidět, že i přes to, že extrémně odlehlé hodnoty signálů bez AF leží v rozmezí signálů s AF, jsou si mediány dat velmi vzdálené a ani kvartily se zde nepřekrývají. Po dopočítání t-testu a wilcoxonova nám dle předpokladu vyšlo zamítnutí nulové hypotézy a přijetí alternativní. Z grafů ROC jde potom vidět, že samotná klasifikace na základě této hodnoty by dokázala v celku přesně rozdělit skupiny pacientů. Hodnota AUC je zde až 0,93. To je 0.1 více než u předchozí nejlepší hodnoty ale pořád je to o 0,5 méně než u celkové klasifikace pomocí neuronové sítě.

Posledním případem je potom rozptyl. Tento rozptyl (podle vzorce 2.6) není počítaný pouze z intervalů RR, ale počítá s body v Poincarého plotu. Pomocí

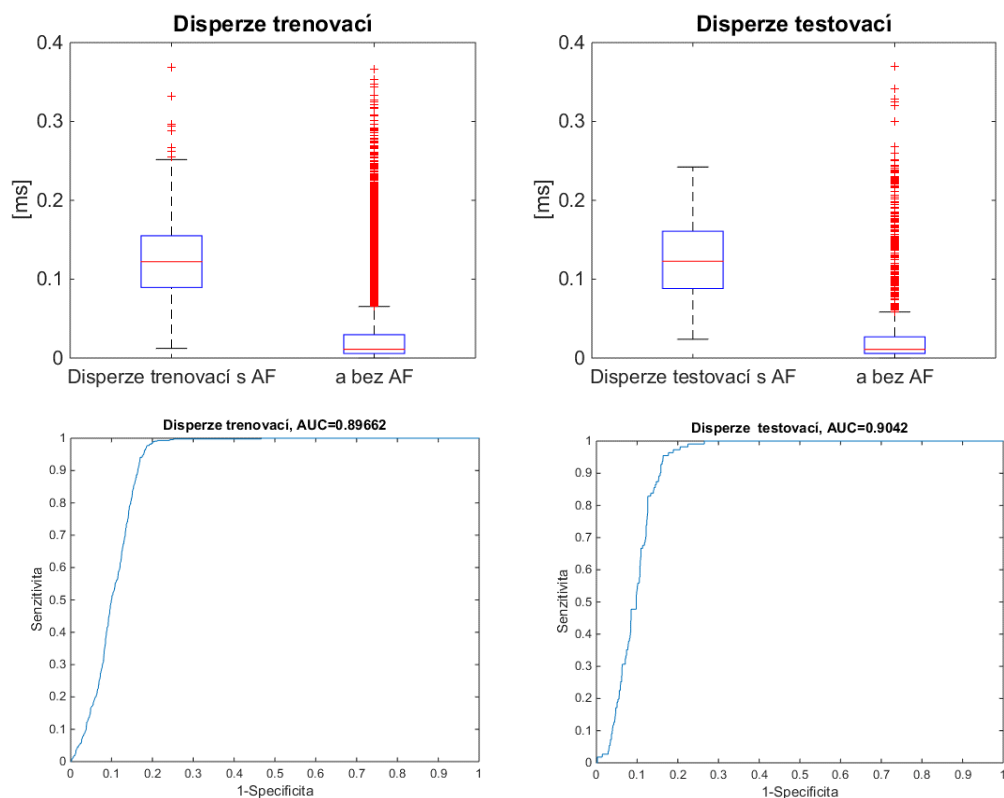
derivace určíme střed grafu a počítáme průměrný rozptyl bodů oproti tomuto středu [5].



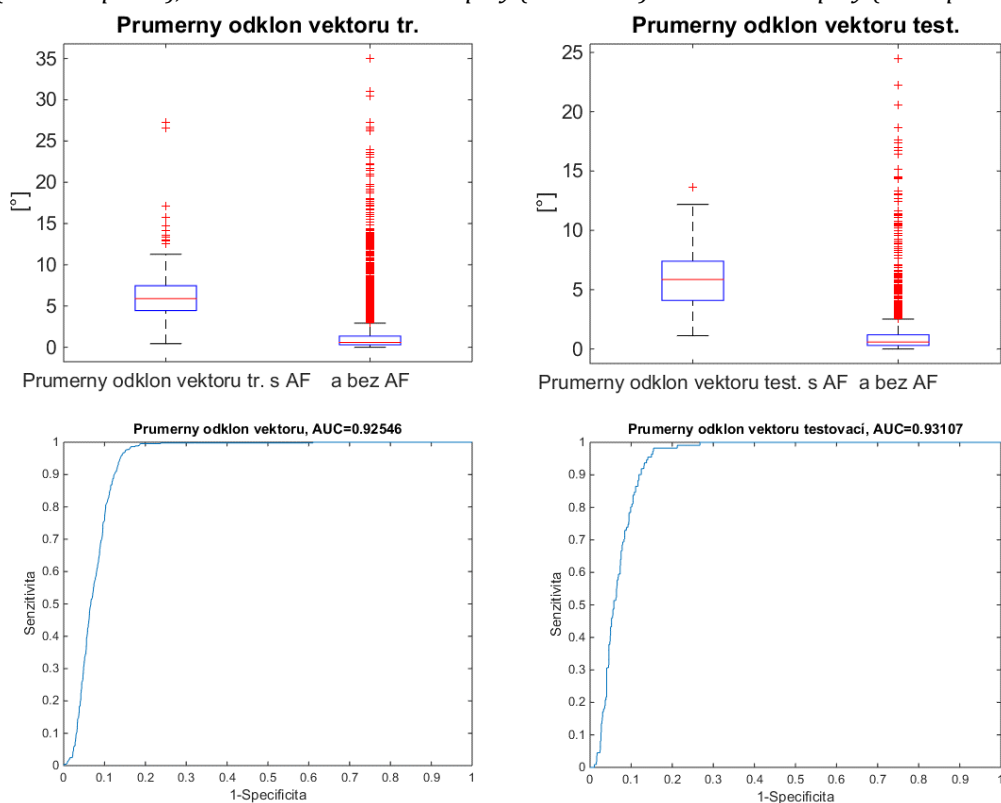
Obr. 11 Grafy krokování (stepping): Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

I tato veličina vychází velmi dobře. I přesto že se grafy rozptylu bodů (Obr. 12) velmi podobají grafům krokování, rozptyl obou skupin je menší než u krokování. Ale hodnoty AUC jsou o málo menší. Jak už jde z grafů vidět, byla i nulová hypotéza o shodě středních hodnot (mediánů) je zamítnuta jak u t-testu tak u testu wilcoxonova. Přijímáme tedy znovu alternativní hypotézu a zařadíme i rozptyl do skupiny pro trénování klasifikačních programů.

Předposlední hodnotou je potom veličina VAI. Tato hodnota se podle grafů na Obr. 13 řadí mezi nejlepší. Z boxplotů jde vyčíst, že medián pacientů s AF leží až nad hranicí 5°, to znamená že více než půlka dat s AF má hodnotu VAI větší než 5° zatímco medián horní kvartil dat bez AF leží pod touto hranicí. To znamená že více než tři čtvrtiny dat s AF jsou pod tímto prahem. Už z tohoto výroku lze předpokládat správnou klasifikaci což jen podkládá hodnota AUC, která je téměř stejná jako u krokování. Tvrzení podkládají znovu i statistické testy, které zamítají nulovou hypotézu a potvrzují tak statistickou významnost i této veličiny

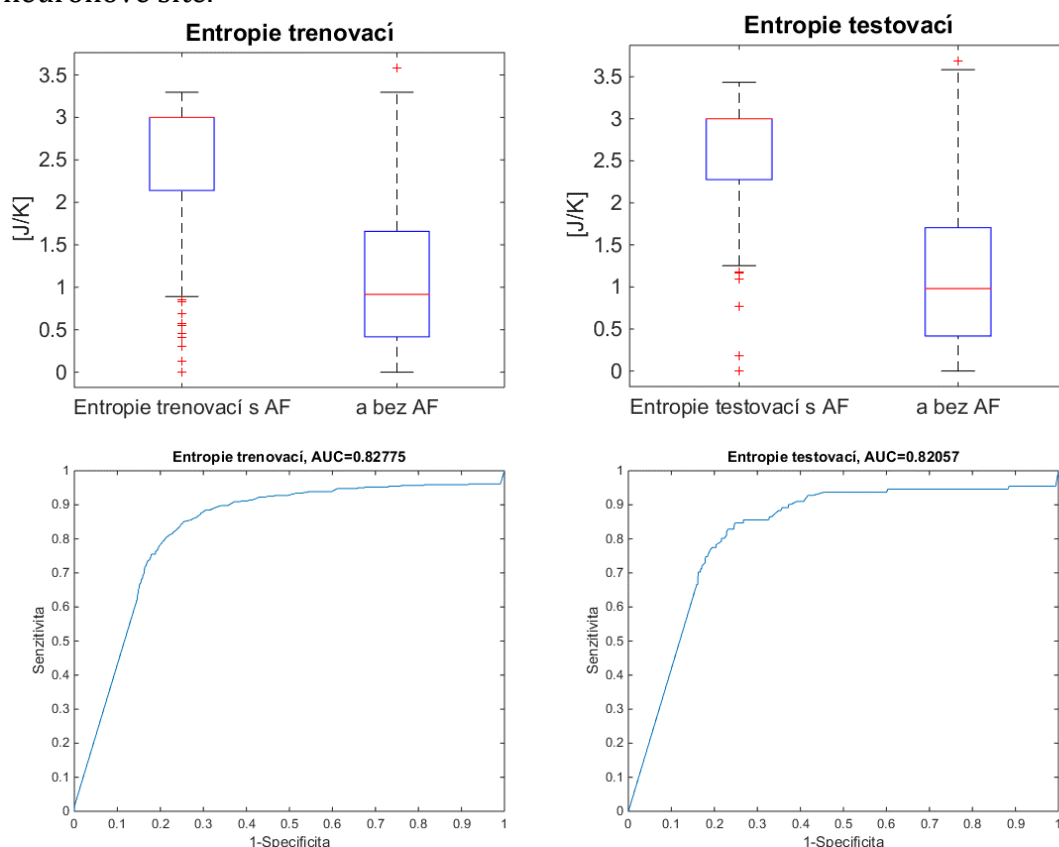


Obr. 12 Grafy rozptylu (dispersion): Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)



Obr. 13 Grafy VAI: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

Poslední veličinou, která vstupuje do výpočtu, je Entropie signálu. I přesto že tato hodnota měla být původně nejlepší v klasifikaci signálů, tak se i přes všechnu snahu nepodařilo doladit její nastavení k hodnotám lepším, než je na Obr. 14. I přesto že boxploty vypadají jako že se obě skupiny zcela překrývají, jdou vidět velmi rozdílné mediány. To znamená že tato veličina sice nerozděluje signál dobře, ale nese jistou informaci. Statistickými testy bylo znovu rozhodnuto o statistické významnosti této veličiny. Proto byla i entropie použita ke klasifikaci pomocí neuronové sítě.



Obr. 14 Grafy Entropie: Boxplot trénovací skupiny (nahore vlevo) a testovací skupiny (nahore vpravo), křivka ROC trénovací skupiny (dole vlevo) a testovací skupiny (dole vpravo)

2.5 Ověření přesnosti a funkce detektoru

K ověření přesnosti detektoru jsou v této práci použita především 4 hlavní kritéria a to sensitivita, specificita, plocha pod křivkou (AUC, z anglického area under curve) a nejefektivnější hodnota reprezentující kvalitu klasifikace je F1 measure.

Sensitivita (2.12) a specificita (2.13) jsou podobně vypočítané veličiny používané pro hodnocení diagnostických testů. Pokud test určuje u osoby přítomnost určitého onemocnění jako je například v této práci AF. Sledujeme potom zda určitá osoba opravdu tímto onemocněním trpí a test ji odhadne jako nemocnou, v top případě řadíme pacienta do skupiny správně pozitivních (TP, true positive). V dalším případě může být osoba zdravá a test ji určí jako zdravou, tehdy se pacient zařadí do skupiny správně negativních (TN, true negative). U zbylých dvou klasifikací se potom jedná o chyby, a to buď chyby typu 1, neboli falešně pozitivní (FP, false positive), u této chyby jde o to, že pacient nemocí netrpí, ale je mu testem predikována. Pacient pak, pokud nebudou udělány jiné testy, které predikci zamítnou, může být léčen na nemoc kterou netrpí. A poslední skupinou jsou potom chyby typu 2, u nichž je pacient, který onemocněním trpí určen jako zdravý, jedná se o výsledek ze skupiny falešně negativních (FN, false negative). V takovém to případě může neléčení nemocného pacienta vést ke zhoršení jeho stavu, až k smrti. Z těchto skupin hodnocení testu se potom dá vypočítat mnoho veličin určujících kvalitu testu. Pro tuto práci je potom nejdůležitější právě senzitivita, což je schopnost testu rozpoznat všechny skutečně nemocné osoby. Jedná se tedy o pravděpodobnost, že test bude u nemocné osoby skutečně pozitivní. A specificita testu, která udává schopnost testu rozeznat všechny zdravé osoby a je to tedy pravděpodobnost, že test bude mít u zdravé osoby negativní výsledek [19].

$$\text{Senzitivita} = \frac{TP}{TP + FN} \quad 2.12$$

$$\text{Specificita} = \frac{TN}{TN + FP} \quad 2.13$$

Plocha pod křivkou je hodnota vypočítaná ze ROC (regression operative curve). Na základě předchozích hodnot nemůžeme hodnotit spojité výstupy. K tomu slouží ROC analýza. Díky této analýze můžeme zjistit jak vhodné veličiny pro samotné rozdělení dat do skupin, tak můžeme pomocí plochy pod touto křivkou hodnotit, jak dobře test diskriminuje dvě skupiny od sebe. V této analýze jde o to, že posunujeme dělicí práh (cut-off) přes všechny hodnoty vstupní proměnné a interpretujeme hodnoty pod prahem jako zdravé jedince a hodnoty nad prahem jako nemocné. Potom pro výsledky sestrojíme graf jako závislost senzitivity na

převrácené hodnotě specificity (FP rate, false positive rate). Pokud by potom klasifikace určila správně pouze polovinu dat (tzn senzitivita a specificita je 0.5) byla by ROC křivka takovéto klasifikace shodná s přímkou pod úhlem 45°. Metoda předpokládá že k takovému výsledku lze dojít právě náhodným výběrem, kdy se jedna polovina určí náhodně jako zdravá a druhá jako nemocná, bez závislosti na vstupních parametrech. Ideální graf dokonalé analýzy by potom vypadal tak, že do určitého prahového bodu by byla senzitivita rovna jedné a hodnota 1 - specificita (FP rate) rovna 0, zatímco od tohoto bodu dále by byla 1 - specificita (FP rate) rovna jedné a senzitivita rovna 0. Takovýto test by potom přesně v prahové hodnotě určil všechny pacienty správně. V praxi nám však musí postačit se takovéto hodnotě co nejvíce přiblížit. Plocha pod takovýmto grafem je přímo rovna 1, přičemž při horší detekci se plocha pod křivkou (AUC) zmenšuje. Tohoto principu využíváme při hodnocení kvality grafu podle plochy pod křivkou. I když tato hodnota není přesně vypovídající pro může být hodnota velmi blízká 1, ale bodu (1,0) se může vzdáleně vyhnout, a navíc je metoda ovlivnitelná nehomogenními daty, kdy jedna skupina je mnohem větší než druhá. To proto, že hodnota AUC nevypovídá přímo o tvaru této křivky.

Proto používáme další kritérium F_1 measure. Toto kritérium je používáno pro kvalifikaci správnosti klasifikace dat. Díky výpočtu tohoto kritéria můžeme porovnat výsledky klasifikace se soutěží „PhysioNet/Computing in Cardiology Challenge 2017” [20]. Zde jsou však data rozdělovány do čtyřech typů, normální signály, AF signály, ostatní rytmy a zašuměné signály, takže hodnoty klasifikace mohou být v soutěži o něco nižší (více skupin znamená více možností špatného zařazení). Pro tuto práci však postačí pouze dělení do dvou skupin, a to signály s AF a ostatní. Většinu zašuměných signálů jsme vyřadily ručně, a zbytek byl vyřazen kritériem podmiňujícím alespoň 4 a více R vln. Pro každou ze skupin potom vypočítáme kritérium podle vzorce 2.14. Kde F_{1a} je hodnota F_1 measure pro klasifikaci atriální fibrilace, ΣA_a je počet všech AF signálů klasifikovaných jako AF (TP), ΣA je počet všech signálů klasifikovaných jako AF a Σa je suma všech signálů predikovaných jako AF. Pro ostatní potom vypočítáme kritérium obdobně. V poslední řadě potom dopočítáme z jednotlivých výpočtu celkovou F_1 measure jako aritmetický průměr těchto dvou hodnot.

$$F_{1a} = \sum \frac{2 * \Sigma A_a}{\Sigma A + \Sigma a} \quad 2.14$$

Na základě těchto čtyř hodnot, především pak právě na základě F_1 , jsme odladily vstupní hodnoty jako je prahová hodnota r pro entropii, množství vnitřních neuronů ve skryté vrstvě neuronové sítě, způsoby rozhodování neuronové sítě, jako je dopředná síť (*feedforwardnet*), *perceptron*, nebo nastavení neuronové sítě pro

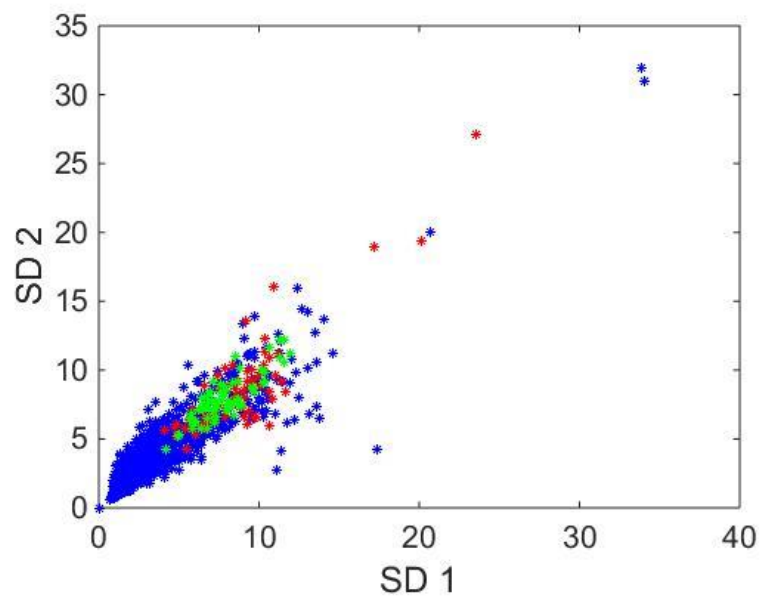
samotné rozpoznávání vzorů (v matlabu funkcí *patternnet*), které je jinak součástí dopředné sítě. Pro odhalení nejlepšího nastavení pro klasifikaci slouží funkce *odladeni.m* v níž jsou pomocí cyklů postupně měněny vstupní veličiny a do tabulky jsou zapisovány výsledky jednotlivých funkcí přesnosti. Posléze je vybrána největší hodnota F_1 measure a na základě její pozice zjištěna ideální kombinace vstupních parametrů pro nejlepší výsledek.

2.6 Sestrojení detektoru

V této části jde o rozřazení dat pomocí k tomu vhodných klasifikací na data, která mají fibrilaci a data která nemají. K tomu můžeme použít tvrdého prahu, rozhodovacího stromu nebo neuronových sítí. Tvrdý práh pro jednotlivé veličiny jsme používali v kapitole 2.4 při sestrojení křivek ROC. Klasifikace podle rozhodovacího stromu v této práci není blíže probírána a to proto, že při jejím zkoušení v programu nebylo dosaženo požadované přesnosti. Ačkoli úspěšnost na trénovací skupině dosahovala až k dokonalému AUC (při přeučení), na skupině testovacích dat vyšla přesnost klasifikace vždy hůře než u neuronových sítí. Proto byly všechny výstupy analýzy, které prošly statistickým testováním použity jako vstupy pro neuronovou síť (v matlabu funkce *train*) pro kterou jsme pomocí cyklu upřesněného v předchozí kapitole vybrali dopřednou neuronovou síť s počtem neuronů ve vnitřní vrstvě rovným 20.

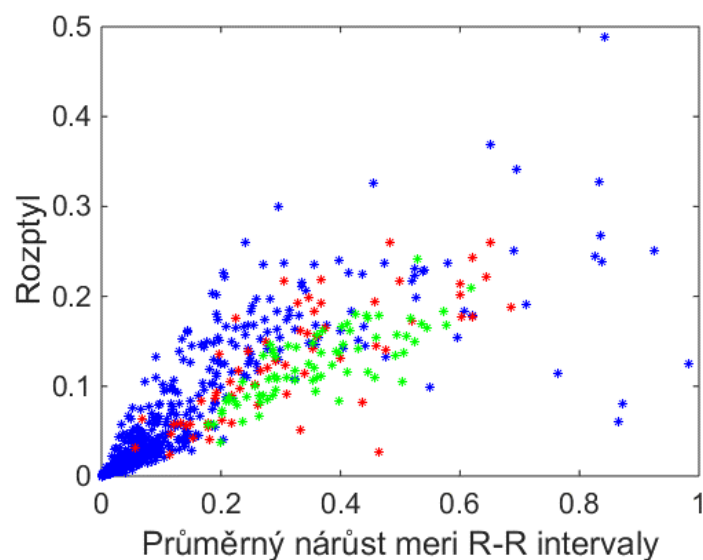
Zbylé nastavení sítě jsme ponechali na výchozím algoritmu Lavenberg-Marquardt (*trainlm*), který je řešen metodou nejmenších čtverců. Optimalizovanou funkcí je zde suma čtverců reziduí, a je kombinací metody největšího gradientu a Gauss-newton. Maximální počet epoch je potom nastaven na 1000. I když se v průběhu učení k takovéto hodnotě ani zdaleka nepřiblížil. [22]

Pro přehled můžeme klasifikovaná data zobrazit i jako grafy, do kterých vyneseme závislosti mezi jednotlivými číselnými reprezentacemi signálu. To proto, aby byla eliminována závislost na pozici vzorku mezi ostatními. Na Obr. 15 můžeme vidět graf závislosti SD 1 a SD 2, červené body jsou pacienti, kteří byli v testovací skupině detekováni špatně. Zelené body jsou pacienti trpící atriální fibrilací detekováni pomocí neuronových sítí správně. A modré body jsou pacienti, kteří jsou zdraví a byly detektorem určeny jako pacienti netrpící atriální fibrilací. Z grafu lze vyčíst, že pod prahovou hodnotou SD 1 i SD 2 = 5 ms leží jen málo pacientů se síňovou fibrilací, nad hodnotou 6,5 ms naopak leží jejich většina. Z Obr. 15 jde potom poznat jaké signály určila metoda neuronových sítí za postižené síňovou fibrilací.



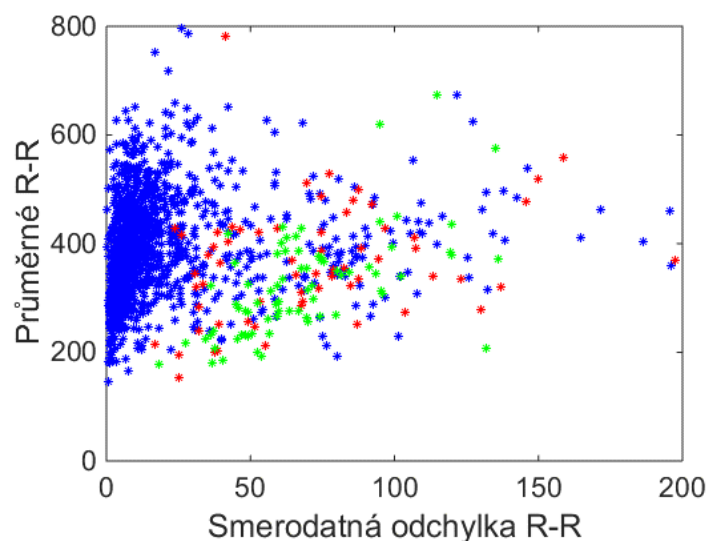
Obr. 15 Závislost Směrodatné odchylky SD 2 [ms] na odchylce SD 1 [ms] detekované doktorem a pomocí funkce neuronových sítí v prostředí matlab (zelené – TP, modré – TN, červené – F)

Dalšími vhodnými grafy jsou pak závislosti pro zobrazení jsou potom závislost rozptylu na krokování [5] (Obr. 16) a závislost průměrné hodnoty RR na směrodatné odchylce RR (Obr. 13). Graf však jen přehledněji znázorňuje to, co lze vyčíst i z box plotu. Důležitější jsou pro nás barevné odlišení, z nichž lze vyčíst, jaké množství bodů je označeno špatně (červené), jaké množství správně jako signály s AF (zelené), a jaké množství správně jako signály bez AF (modré).



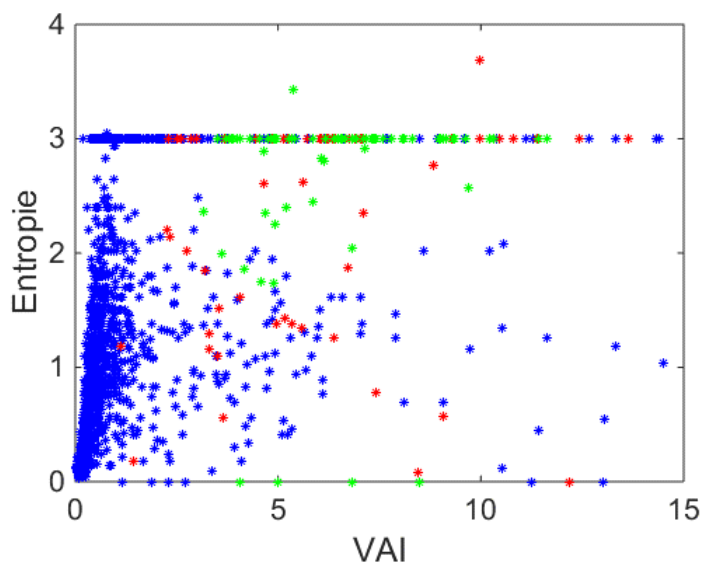
Obr. 16 Závislost rozptylu [ms] na krokování [ms] detekované doktorem a pomocí funkce neuronových sítí v prostředí matlab (zelené – TP, modré – TN, červené – F)

Dalším grafem pro znázornění je graf závislosti průměrné hodnoty RR na směrodatné odchylce (Obr. 17).



Obr. 17 Závislost průměrného RR [ms] na směrodatné odchylce RR [ms] detekované doktorem a pomocí funkce neuronových sítí v prostředí matlab (zelené – TP, modré – TN, červené – F)

A posledním grafem je pak závislost entropie na VAI (Obr. 18). Entropie však dosahuje kvůli výpočtu v programu *matlab* hodnot Inf (nekonečna) pro poměr shod 1/0, NaN (neurčeno) pro poměr shod 0/0 a $-Inf$ (-nekonečno) pro poměr shod 0/1. Tyto hodnoty jsou ověřeny vnitřní funkcí při výpočtu entropie a jsou nastaveny tak aby odpovídaly realitě, ale zanáší do grafu určité zkreslení.



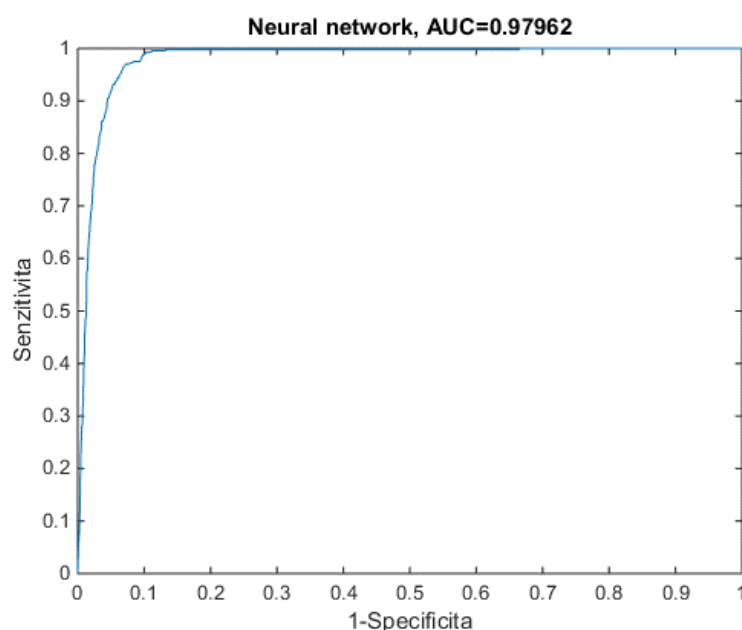
Obr. 18 Závislost Entropie [J/K] na VAI [°] detekované doktorem a pomocí funkce neuronových sítí v prostředí matlab (zelené – TP, modré – TN, červené – F)

Na těchto grafech můžeme vidět závislost positivity či negativity AF na pozici, respektive na jednotlivých hodnotách, v grafu. Důležité je, že vstupem do neuronové sítě je všech 9 původních hodnot pro každý signál, v její skryté vrstvě je potom 20 neuronů a výstupem je hodnota v rozmezí od 0 do 1. Z celkem 6884 signálů je 4991 signálů bez atriální fibrilace a 442 signálů s atriální fibrilací vybráno jako trénovací skupina. Signály na trénování a testování jsou vybrány předem. Pro lepší trénování neuronových sítí je vektor signálů s atriální fibrilací zkopírován 10krát za sebou, aby byly obě skupiny přibližně stejně dlouhé a následně jsou náhodně promíchány. Do testovací skupiny bylo potom vybráno 1247 signálů bez AF a 111 s AF. V grafech jsou potom znázorněny trénovací i testovací skupiny, to proto, aby bylo vidět že není příliš důležité, jak se rozdělí trénovací skupina, ale především testovací, protože na tu není neuronová síť naučená.

3 ZÁVĚR

Nejprve jsme graficky a statistickou metodou nepárového t-testu a ověřily správnost Jednotlivých klasifikací (SD 1, SD 2, plocha elipsy, krokování, rozptyl, střední hodnota RR intervalů, směrodatná odchylka RR intervalů, VAI a entropie). Přičemž všechny tyto veličiny se ukázaly jako statisticky významné pro detekci síňových fibrilací na hladině významnosti $\alpha=0,05$ %.

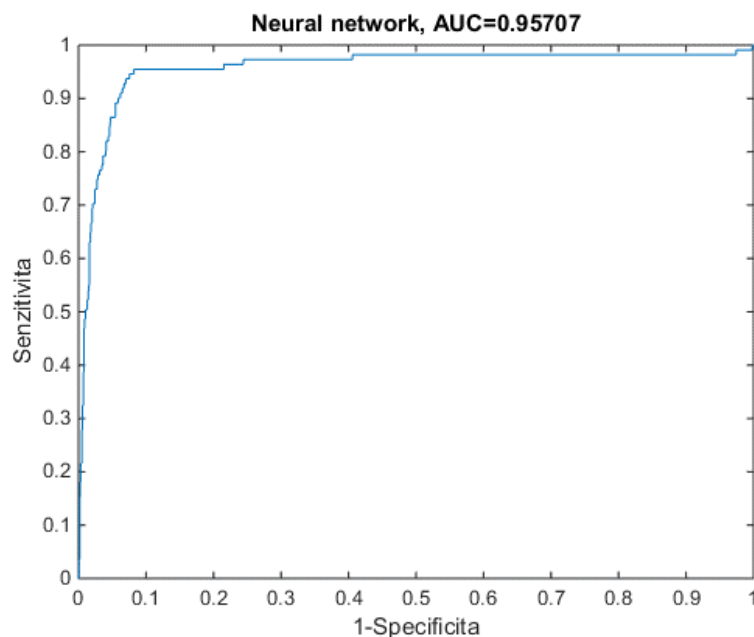
Dále je provedena detekce pomocí dopředné neuronové sítě, na základě níž jsou data z testovací skupiny rozděleny na dvě skupiny. První skupinou jsou data, u kterých jsme detekovali AF a druhou data u kterých ne. Tato metoda by se dala nahradit například tvrdým prahem nebo rozhodovacím stromem.



Obr. 19 ROC trénovací skupiny dat (AUC = 0.980)

S neuronovou sítí jsme dosáhli senzitivity 74,55 % a specifity 97,11 %. To znamená že pouze 75 z 100 pacientů trpících síňovou fibrilací je označeno za nemocné, a 97 z 100 zdravých pacientů je označeno jako zdravý. To ale znamená že 25 pacientů ze 100 by nebyla nemoc odhalena a 3 pacienti z 100 by byla určena nemoc kterou netrpí. ROC z Obr. 19 potom ukazuje že plocha pod křivkou je pro trénovací množinu vysoká (AUC = 0,980) a pro testovací množinu (Obr. 20) nižší (AUC = 0,957). Takováto klasifikace není příliš dobrá, avšak je méně chybová než jiné nastavení. Pomocí rozhodovacího prahu se dá výstup neuronové sítě nastavit i jinak. Ale počet celkových chyb se tím pouze zvýší (snížením TN se zvýší FN a naopak). Nevýhodou učení neuronových sítí je to, že se jedná částečně o náhodný proces, při kterém se mohou výsledky učení lišit i pokud se vstup nijak nezmění. Důvodem je že neuronové sítě patří mezi stochastické procesy, které jsou velmi

citlivé na své vstupní nastavení. Dochází u nich k náhodnému prohledávání prostoru, a proto nemusíme při špatném nastavení na řešení narazit. Výsledkem se chyby klasifikace při opakovaných pokusech pohybují mezi 55 až 66. Chyby 2. typu (FN) jsou závažnější, a pohybují se v rozmezí od 26 až 43 chybami. Což je velká část ze 111 celkových signálů. Chyby 1. typu (FP) se pak pohybují v rozmezí od 15 do 40 chyb. Jak jde vidět v součtu je chyb vždy podobně, takže čím více chyb 1. typu tím méně chyb 2. typu a naopak. Pro nás je však lepší více chyb 1. typu, navíc jejich poměr proti celkovým 1247 signálům bude vždy menší.



Obr. 20 ROC testovací skupiny dat (AUC = 0.957)

Při porovnání s výsledky soutěže „PhysioNet/Computing in Cardiology Challenge 2017” [20] bylo dosaženo srovnatelných výsledků. Uvedených v tabulce:

TVŮRCE	F ₁ TESTOVACÍ	F ₁ TRÉNOVACÍ
TATO PRÁCE	0,852	0,9461
TEIJEIRO A SPOL.	0,831	0,990
DATTA A SPOL.	0,826	0,968
ZAHIBI A SPOL.	0,825	0,990
HONG A SPOL.	0,822	0,859

Chybou při srovnání může být už dříve zmíněné rozdělování do skupin. Chyba při nastavování vstupních parametrů jak neuronové sítě, tak například i výpočtu entropie.

LITERATURA

- [1] ŠTEJFA, M. ŠUMBERA, J. BRAVENÝ, P. Základy elektrokardiografie, Brno: Masarykova univerzita, fakulta lékařská, 1991, 147 s. ISBN 80-210-0265-4.
- [2] AUTOR NEZNÁMÝ. Stavba srdce – obrázek. Třídní stránky. [online]. 29.11.2014 [cit. 2014-11-29]. Dostupné z: <http://gymplmb.blogspot.cz/2013/01/obrazek.html>
- [3] BRAGARD, J. Universidad de Navarra [online]. 2010 [cit. 2011-05-24]. Analysis of time series of physiological data. Upravené a dostupné z: <http://fisica.unav.es/~jbragard/detailecg.gif>
- [4] KOZUMPLÍK J.: Multitaktní systémy. Elektronické skriptum UBMI FEKT VUTBR 2005, dostupné z: https://www.feec.vutbr.cz/et/skripta/ubmi/Multitaktni_systemy_S.pdf
- [5] TAYEL, Mazhar B. a Eslam I. ALSABA. Poincaré Plot for Heart Rate Variability. International Journal of Biomedical and Biological Engineering [online]. World Academy of Science, Engineering and Technology, 2015, 2015, 9(9), 4 [cit. 2018-01-04]. DOI: urn:dai:10.1999/1307-6892/10002615. Dostupné z: <https://waset.org/publication/Poincaré-Plot-for-Heart-Rate-Variability/10002615>
- [6] JINHO, Park, Lee SANGWOOK a Jeon MOONGU. Atrial fibrillation detection by heart rate variability in Poincare plot. BioMedical Engineering OnLine [online]. © Park et al; licensee BioMed Central Ltd. 2009, 2009, 11 prosinec 2009, 2009(8:38), 12 [cit. 2018-01-04]. Dostupné z: <https://doi.org/10.1186/1475-925X-8-38>
- [7] PISKORSKI, Jarosław a Przemysław GUZIK. Filtering Poincaré plots. COMPUTATIONAL METHODS IN SCIENCE AND TECHNOLOGY [online]. Institute of Physics, University of Zielona Góra, Poland; Department of Cardiology – Intensive Therapy, University of Medical Sciences in Poznań, Poznań, Poland, 2005, 11(1)(39-48(2005), 10 [cit. 2018-01-04]. DOI: DOI:10.12921/cmst.2005.11.01.39-48. Dostupné z: http://cmst.eu/wp-content/uploads/files/10.12921_cmst.2005.11.01.39-48_Piskorski.pdf

- [8] Ruan, X., Liu, C., Liu, C., Wang, X., & Li, P. (2011). Automatic detection of atrial fibrillation using R-R interval signal. 2011 4th International Conference on Biomedical Engineering and Informatics (BMEI), 644–647. Dostupné z: https://www.researchgate.net/publication/220715249_Automatic_detection_of_a_trial_fibrillation_using_R-R_interval_signal
- [9] Co to je entropie. Wwww.tzb-info.cz [online]. 2002 [cit. 2019-11-04]. Dostupné z: <http://www.tzb-info.cz/925-co-to-je-entropie>
- [10] Neznámý. (n.d.). 2. Teorie informace 2.1 Vznik a vývoj teorie informace. Dostupné z: http://www.uai.fme.vutbr.cz/~matousek/TIK/dokumenty/osmera_kap2.pdf
- [11] Travis J Moss, D. E. L., & Moorman, J. R. (2014). Physiological Measurement Local dynamics of heart rate: detection and prognostic implications Related content. Dostupné z: <https://pdfs.semanticscholar.org/426a/19f2871ba3fe5f2f3d427491f16330be59b2.pdf>
- [12] Lake, D. E., & Moorman, J. R. (2011). Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. American Journal of Physiology-Heart and Circulatory Physiology, 300(1), H319–H325. Dostupné z: https://www.researchgate.net/publication/47635928_Accurate_estimation_of_entropy_in_very_short_physiological_time_series_The_problem_of_atrial_fibrillation_detection_in_implanted_ventricular_devices/download
- [14] Alcaraz, R., Abásolo, D., Hornero, R., & Rieta, J. J. (2010). Optimal parameters study for sample entropy-based atrial fibrillation organization analysis. Computer Methods and Programs in Biomedicine, 99(1), 124–132. Dostupné z: <https://vdocuments.site/optimal-parameters-study-for-sample-entropy-based-atrial-fibrillation-organization.html>
- [15] Holčík, J., Analýza a klasifikace signálů. Nakladatelství VUT v Brně, Brno (1992)
- [16] Kotek, Z., Brůha, I., Chalupa, V., Jelínek, J., Adaptivní a učící se systémy. SNTL, Praha (1980)
- [17] Pavlík, J. (2005). Aplikovaná statistika, 173, Retrieved from http://147.33.74.135/knihy/uid_isbn-80-7080-569-2/pages-img/001.html
- [18] Veterinární a farmaceutická fakulta, B. (n.d.). Parametrické testy - Studentův t-test. Použito 14. dubna 2019, z teorie, přednáška 3, Dostupné z: <https://cit.vfu.cz/statpotr/POTR/Teorie/Predn3/ttest.htm>
- [19] Masarykova univerzita, B. (2014). Analýza dat pro Neurovědy. Dostupné z: https://is.muni.cz/el/1411/jaro2014/DSAN01/um/45910621/Analiza_dat_pro_Neurovedy_-_blok_6.pdf

- [20] Clifford, G. D., Liu, C., Moody, B., Lehman, L. H., Silva, I., Li, Q., ... Mark, R. G. (2017). AF Classification from a Short Single Lead ECG Recording: the PhysioNet/Computing in Cardiology Challenge 2017. Computing in Cardiology, 44. Dostupné z: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5978770/>
- [21] MATLAB - MathWorks - MATLAB & Simulink. (n.d.).[Internet] Použito Květen 8, 2019, z <https://uk.mathworks.com/products/matlab.html>
- [22] Wolfram Mathematica: Documentation center, [Internet], <http://reference.wolfram.com/applications/neuralnetworks/NeuralNetworkTheory/2.5.3.html>

Seznam symbolů, veličin a zkratek

AF	-	Síňová (atriální) fibrilace
AUC	-	Plocha pod křivkou (Area under curve)
EKG	-	Elektrokardiogram (Elektrokardiograf)
FEKT	-	Fakulta elektrotechniky a komunikačních technologií
FP	-	Špatně pozitivní (False Positive)
FN	-	Špatně negativní (False Negative)
VUT	-	Vysoké učení technické v Brně
R-R	-	Interval od R vlny k R vlně
ROC	-	Regresivní operativní křivka (Regresive operative curve)
SD	-	Směrodatná odchylka (Standart Deviation)
TP	-	Správně pozitivní (True Positive)
TN	-	Správně negativní (True Negative)

Seznam příloh

Příloha 1. CD

Obsah CD:

- Project_BTL.m
- rozlisovaci_schopnost.m
- boxploty.m
- sampleEntropy.m
- odladeni.m
- grsSave.m

- BTL_qrs.mat
- indAfibTrainTest.mat
- mainCategories.mat